TESIS DOCTORAL

Juegos Markovianos Discretos. Una Aproximación a Modelos de Desarrollo Sostenible

Autor: Omar Jesús Casas López

Directora: Dra. Dña. Rosario Romera Ayllón

Doctorado en Ingeniería Matemática

Departamento de Estadística

UNIVERSIDAD CARLOS III DE MADRID

Leganés, Enero 2010

A mi hijo Andy

Vi un alto muro y como tenía la premonición de un enigma, algo que podría estar escondido detrás del muro, trepé por él con alguna dificultad. Sin embargo, al otro lado caí en una selva y tuve que abrirme camino con gran esfuerzo hasta que llegué a la puerta abierta, la puerta abierta de las matemáticas. A partir de aquí, caminos muy transitados conducían en todas las direcciones y desde entonces he pasado tiempo allí. A veces pienso que ya he recorrido todo el área, que ya he pisado todos los caminos y admirado todas las vistas, y entonces descubro de repente un nuevo camino y experimento nuevas delicias.

Maurits Cornelis Escher

Agradecimientos

"... se hace camino al andar"

Antonio Machado

Por fin después de algunos años de andar por este camino, con una gran variedad de obstáculos de toda índole, estoy llegando a la meta del mismo, que no es más que otro punto de partida. En este trayecto he encontrado muchas y buenas experiencias, gracias a todos los que me han acompañado en esta etapa.

Durante estos años, muchas personas han intervenido de uno u otro modo en la realización de este proyecto, sin las cuales no habría sido posible lograr llegar a su final, por ello quiero mostrarles mi gratitud, pero son tantas que si las menciono a todas entonces esta sería la parte más larga de la tesis, sólo mencionaré a los implicados más directos y si alguno se me olvida, estoy seguro que me sabrá disculpar.

En primer lugar, quisiera agradecer enormemente a mi directora de tesis Charo por abrirme las puertas para hacer el doctorado, por todo lo que me ha enseñado, el tiempo que me ha dedicado, la confianza que ha depositado en mí, aunque yo mismo no estuviera tan seguro, y sobre todo las ideas matemáticas en las que me ha propuesto trabajar. Además de ser la guía y orientación de esta memoria, también se ha implicado directamente constituyendo una ayuda tanto desde el punto de vista profesional como humano, aconsejándome al tomar decisiones, mostrando mucha paciencia y optimismo en los momentos más difíciles, motivándome a continuar adelante, etc. Muchas gracias por guiarme en la perspectiva estocástica de la optimización en Teoría de Juegos, así como por haber convertido las dificultades del Control Óptimo y los Procesos Estocásticos en cuestiones asequibles con explicaciones intuitivas, por darme la oportunidad de conocer y aprender de diferentes investigadores del área, a través de congresos, seminarios y cursos, y por haber realizado una labor minuciosa de la revisión de todos los contenidos de esta memoria, así como de los working papers y artículos derivados de la misma.

Agradezco los comentarios y sugerencias sobre mi trabajo, que me han hecho en diversas ocasiones y lugares, Elisenda Molina, Javier Nogales, Maria Albareda, Mike Wiper, Henry Tulkens, Onésimo Hernández-Lerma, Michael Finus, Emilio Cerdá, Guiomar Martín-Herrán, Jean-Marie Grether y Juan Pablo Rincón-Zapatero.

Quiero mostrar mi gratitud al Ministerio de Educación y Ciencia y a la sección de educación y cultura de la Comunidad de Madrid, por el apoyo financiero ofrecido mediante los proyectos de investigación SEJ2004-03303 (2004/02934/001 y 002) y SEJ2007-64500 (2007/04438/001) del MEC, y UC3M-ECO-05-061 (2006/03563/001) y CCG06-UC3M/HUM-0866 (2007/04081/001) de la CAM. En particular quiero agradecer a su investigador principal, Daniel Peña, por haberme permitido participar en estos proyectos.

Agradezco al Departamento de Estadística, a su director Javier Prieto, a la organización del Doctorado en Ingeniería Matemática, a su director Ismael Sánchez, y a la Fundación Carlos III que me han otorgado una beca para poder terminar este proyecto. También, quiero dar las gracias a todos los miembros del departamento, en especial al personal de secretaría Lidia, Susana y Paco, por ayudarme amablemente en los múltiples y ásperos trámites de papeles, entre otras cosas.

A mis compañeros asociados Blanca, Asun, Redchuk y María Felix, que han sido un apoyo fundamental, ofreciéndome siempre su ayuda, tanto en la investigación como en la docencia, además de ser amigos con los que he disfrutado de muy buenos ratos en comidas y cafés.

Deseo manifestar mi agradecimiento a aquellos con los que he compartido despacho en estos años Aurelio, Carmen, David y Bárbara, por soportar mis cambios de humor pacientemente. Además a mis compañeros de doctorado, a los que ya han logrado terminar, por compartir conmigo sus experiencias Alba, Aurora, Conchi, Pepa, Santi, Peter, Oscar y Sara; a los que siguen en el camino Ester, Alex, José Antonio, Maye, Adolfo, Kenedy, Andrea, Ana Laura, Nuria, Henry, Dalia, Angel y Hector; a los que fueron mis alumnos de la Licenciatura y hemos sido compañeros de doctorado Javier, Lee y Miguel Angel, y a mis colegas más cercanos del Departamento con los que he compartido muy buenos ratos y comentarios de comida y café Helena, Sergio, Juan Miguel, Andrés, Amparo, Mayte, Ana, Isabel, Alfredo, Roland, Estefano, Alberto, Aurea, Belén, entre otros; a todos ellos les agradezco por haber escuchado y contestado la mayoría de mis preguntas sobre Latex y Matlab, en gran parte gracias a ellos, esta memoria presenta el aspecto actual.

A mis amigas Aurea Grané y Belén Martín-Barragán les agradezco especialmente por la valiosa ayuda que me han ofrecido desinteresadamente con los códigos de los problemas planteados.

En resumen, todos mis compañeros del Departamento han hecho que venir a trabajar, contra todo pronóstico, me resulte agradable.

Quiero expresar mi agradecimiento a mis profesoras de inglés Anne Merry y Sarah Dague por la ayuda que me han prestado en la preparación de presentaciones.

También recuerdo en este momento a los colegas del Departamento de Matemáticas de la Universidad Autónoma de Madrid, Dani, Chema, Javier, Pablo, Mayte, Ana, Raúl, Susi, que hace algunos años me acogieron como uno más de la pandilla, me enseñaron los primeros pasos de Latex y de búsqueda en internet, y me contagiaron con el virus de la investigación matemática. Durante los cursos tuve la suerte de compartir con Cristina, Annia, Ester, Rocío entre otros y Antonio Cuevas, Ana Justel, José Ramón, Adolfo, Santiago, Josechu, etc, como profesores.

Quiero dar las gracias también a varios amigos que han influido de formas diferentes en esta tesis, a mis amigos de la Habana Alibeit, Helmo y Zayda, a los más cercanos Charo, Gustavo, Alma, Orestes, Mirna, Miguel, Jorge, Adolfo, y otros, que se han interesado por mí, preocupándose por mi trabajo, por mi estado de ánimo,... etc., alentándome en los momentos más difíciles y disfrutando conmigo de las buenas etapas. Muchas gracias de todo corazón, sin vosotros no hubiese sido posible llevar a cabo este trabajo.

Quiero agradecer a mis padres que me enseñaron a andar, a mis hermanos Jesús, Zoe y Mario que desde el otro lado del Atlántico se han preocupado y me han motivado a seguir adelante, a Rolf que ha sido un gran apoyo aliviándome considerablemente el peso de esta parte del camino, y he dedicado esta memoria a mi hijo Andrés ya que es el motor que me impulsa a seguir andando.

Resumen

En los últimos años ha aumentado la preocupación por la contaminación ambiental y su relación con el cambio climático. Las Naciones Unidas han promovido numerosas reuniones y cumbres de Jefes de Estados, para intentar llegar a un acuerdo entre todos los países con el objetivo de disminuir las emisiones de CO_2 a la atmósfera. Desde el Protocolo de Kyoto en 1997 hasta la conferencia de Copenhague en diciembre del 2009, ponen de manifiesto esta preocupación, así como las dificultades que impiden alcanzar soluciones satisfactorias para todos los países y que a la vez consigan el objetivo de reducción de las emisiones.

El contexto teórico en el que estos problemas vienen formulándose desde los años 70 es la Teoría de Juegos, inicialmente como juegos estáticos y posteriormente como juegos dinámicos. Hasta los años finales del siglo XX, la mayor parte de los resultados se han obtenido para modelos deterministas, y es en esta primera década del siglo XXI cuando se comienzan a abordar formulaciones estocásticas para problemas particulares en este ámbito de modelos de desarrollo sostenible.

El objetivo de esta tesis es proporcionar modelos estocásticos para el control del stock acumulado de contaminación ambiental, formulados como Procesos de Decisión de Markov (MDP) con horizonte finito. En este contexto, un importante paradigma es la minimización del funcional de coste que depende de la evolución del stock de contaminación (sistema) afectado por perturbaciones aleatorias, a lo largo de un horizonte finito de T etapas, esto es, el problema tipo llamado TSO (T-stage stochastic optimization problem).

El paradigma del problema tipo TSO es suficientemente general como para poder ser tomado como base de solución de problemas de horizonte infinito, por ejemplo mediante la aplicación de técnicas tipo horizonte en retroceso (receding horizon) o de forma directa mediante algoritmos de iteración de valores.

Los escenarios no cooperativos que en Teoría de Juegos proporcionan los equilibrios de

Nash, son el punto de partida para los distintos problemas TSO que se abordan en esta tesis.

El planteamiento y la solución de los problemas que se formulan al asumir comportamientos cooperativo y no cooperativo de los jugadores (países) o decisores, así como la posibilidad de utilizar alicientes monetarios, llamados transferencias, para incentivar la cooperación, son los resultados presentados en los Capítulos 2 y 3 de esta tesis. En la idea de proporcionar a cada país estrategias que le permitan tomar decisiones que se aproximen al óptimo internacional, esto es al óptimo cooperativo, se formula un novedoso criterio probabilístico de optimización, que difiere del habitual basado en la minimización del valor esperado de un cierto funcional de coste, que es el estandar para problemas Markovianos de decisión. Los problemas de optimización derivados de este planteamiento y sus resultados se presentan en el Capítulo 4.

Para ilustrar la modelización de los problemas de optimización estocástica abordados en la tesis, se ha utilizado un mismo ejemplo numérico basado en datos reales, que presenta seis países o regiones como jugadores con un horizonte de 40 años y cuyos datos iniciales se han tomado en 1990, base de la comparación del Protocolo de Kyoto.

Es bien conocido que para los MDP, su solución analítica puede hallarse, sólo en algunos casos sencillos (por ejemplo en sistemas lineales con coste cuadrático). La metodología más habitual para aproximar la solución desarrollando técnicas numéricas con un coste computacional razonable, se obtienen mediante técnicas de Programación Dinámica (aproximada), y esta es la metodología utilizada en este trabajo.

Los códigos desarrollados para la obtención de los escenarios aleatorios que han permitido evaluar las soluciones de ciclo cerrado, como corresponde a problemas de decisión markovianos, han sido específicamente obtenidos en Matlab para cada uno de los problemas estudiados en esta tesis.

Contents

List of Figures

List of Tables

Capítulo 1

Antecedentes y Motivación

"What's the good of Mercator's
North Poles and Equators,
Tropics, Zones and Meridians Lines?"
So the Bellman would cry:
And the crew would reply
"They are merely conventional signs!"

Lewis Carroll

Resumen

Este Capítulo 1 contiene una revisión del material teórico necesario para desarrollar los capítulos posteriores. En particular, se introducen la descripción y los resultados básicos sobre Juegos Markovianos y la existencia de equilibrios, además se presenta el concepto de Desarrollo Sostenible y se comentan algunos modelos que han aparecido en la literatura relacionada con sostenibilidad, de diferentes áreas, en los cuales se aplican técnicas de Teoría de Juegos o Programación Dinámica en su resolución. Al final de este capítulo se comenta la estructura y objetivos de la esta memoria.

1.1 Introducción

A lo largo de las secciones introductorias de los Capítulos 2, 3 y 4 de esta memoria se van a presentar los antecedentes y resultados aparecidos en la literatura especializada, en relación con cada uno de los modelos estocásticos, procesos de decisión de Markov (MDP), propuestos en el correspondiente Capítulo y los problemas de ellos derivados, problemas de optimización dinámica estocástica en tiempo discreto y con horizonte finito. En este Capítulo se van a recoger los elementos teóricos y su conexión con los problemas de control del stock de polución ambiental como parte de las conexiones entre Juegos Markovianos (MG) y modelos de Desarrollo Sostenible.

En la Sección 1.2 de este capítulo se describe la relación entre la Programación Dinámica (DP) y los Juegos Markovianos (MG), y se estudia el problema de la existencia de equilibrios en Juegos Markovianos, enfocando sobre la técnica precisa y los problemas conceptuales que forman la base para obtener condiciones de existencia satisfactorias y generales. El análisis incluye Juegos Markovianos generales. Además se revisan algunos tópicos importantes, tales como los conceptos de estrategia y de equilibrio perfecto en subjuego. También se presenta un análisis de la existencia de equilibrio perfecto en subjuego, en particular sobre una clase de estrategias llamadas estrategias Markovianas. Esta Sección 1.2 está dividido en cuatro partes, en la Subsección 1.2.1 se define un Juego Markoviano General. En la Subsección 1.2.2, se formaliza la noción de estrategia como un plan contingente de acción y describe un equilibrio del Juego Markoviano, también se introducen los conceptos mas importantes de estrategia y de Equilibrio Perfecto de Markov (MPE). La Subsección 1.2.3 aporta un comentario sobre la generalidad de la estructura descrita. Finalmente la Subsección 1.2.4 trata sobre el problema de la existencia de equilibrios, en particular del Equilibrio Perfecto de Markov (MPE).

En la Sección 1.3 se presenta la definición de Desarrollo Sostenible y se denotan los elementos del mismo en términos de Teoría de Juegos. Además se comentan, a modo de ejemplos, algunos trabajos de diferentes campos que desarrollan modelos de Desarrollo Sostenible aplicando técnicas de Teoría de Juegos.

En la Sección 1.4 se presenta la estructura de la tesis, se proponen las líneas de investigación a seguir y los objetivos que se pretenden lograr con el presente trabajo.

1.2 Juegos Markovianos

Los Juegos Markovianos o estocásticos comenzaron a tener mucha popularidad, desde los años 70, para el análisis de interacción estratégica en el contexto de la economía dinámica. Este desarrollo se obtiene por la fusión de los Problemas de Programación Dinámica Estacionaria y los Juegos Repetidos. Como en los Problemas de Programación Dinámica Estacionaria, los Juegos Markovianos plantean la existencia de una variable *estado* definida para capturar el ambiente del juego en cada período de tiempo, pero que se mueve a través del tiempo en respuesta a las acciones tomadas en el juego. Como en los Juegos Repetidos, los Juegos Markovianos permiten la existencia de múltiples decisores o jugadores en el modelo. Referencias básicas para este tipo de juego se pueden encontrar en Dutta (2001), Amir (2003), Neyman and Sorin (2003), Powell (2007), entre otros.

Esta fusión permite una rica variedad de posibilidades en Juegos Markovianos, por ejemplo, una acción de un jugador puede afectar sus futuros resultados de dos maneras:

1. Mediante el efecto que tienen sobre el medio físico en el cual se pueden tomar decisiones futuras.

2. Mediante el impacto sobre el comportamiento de otros jugadores en el modelo.

Los Problemas de Programación Dinámica Estacionaria permiten lo primero, debido a que son problemas unipersonales, los Juegos Repetidos sólo permiten lo segundo, debido a que el escenario del juego debe ser inalterable. Como consecuencia de estas características, el Juego Markoviano es lo más apropiado para el análisis de modelos en los cuales es necesario permitir un cambio tomando decisiones de medio y competición imperfecta.

El entorno de Juegos Markovianos ha visto múltiples aplicaciones en los últimos años, entre otras:

- estrategias de inversión bajo oligopolios (Funderberg and Tirole (1983), Davies and Liebman (2006), Calzolari and Lambertini (2007), Cellini and Lambertini (2007), Figuieres (2009));

- investigación y desarrollo bajo competición imperfecta (Hunt (2006), Cordes (2008), Reksulak et al. (2008));

- modelos dinámicos de oligopolio (Krakel and Sliwka (2006), Matsumoto and Serizawa (2007), McCausland (2007), Bischi et al. (2007));

- desarrollo económico como consecuencia de sistemas heredados (Blackburn and Cipriani (2005), Nowak (2006), Ibragimov (2008));.

- juegos asociados (Radner et al. (1986), Clementi and Hopenhayn (2006)) y problemas de agente principal (Dutta and Radner (1999), MacLeod (2003)).

La aproximación entre la Programación Dinámica y los Juegos Markovianos ha jugado un papel central en recientes avances concernientes a Juegos Repetidos, lo mismo bajo información completa que bajo información incompleta. A pesar de estos avances, hay que admitir que el uso de Juegos Markovianos en la literatura ha sido limitado considerando la generalidad del marco. La razón se debe a que los Juegos Markovianos son significativamente más complejos desde un punto de vista técnico que los Juegos Repetidos o la Programación Dinámica Estacionaria. El equilibrio en un Juego Markoviano muy simple puede tomar formas muy complejas y nada intuitivo.

La existencia de un equilibrio perfecto en subjuego (subgame perfect) en Juegos Markovianos generales ha sido establecida bajo algunas restricciones. Los resultados disponibles no son satisfactorios en al menos dos terrenos:

1. Todos los resultados generales sobre la existencia imponen requerimientos técnicos sobre el mecanismo de transición para las variables estado que impiden que la transición pueda ser *determinista*.

2. Asumiendo un entorno Markoviano *estacionario*, no hay resultados que demuestren la existencia de equilibrios en estrategias Markovianas, incluso con las restricciones sobre el mecanismo de transición antes mencionadas.

La ausencia de un resultado que permita transiciones deterministas es una seria carencia en la literatura, ya que estos mecanismos son empleados en una gran cantidad de aplicaciones. Concerniente al segundo punto, entre los mejores resultados obtenidos está el de Mertens and Parthasarathy (1991) sobre la existencia de un equilibrio perfecto en subjuego con estrategias posiblemente no Markovianas bajo condiciones muy generales.

Los resultados de Mertens and Parthasarathy (1991) en Juegos Markovianos, representan las condiciones más generales bajo las cuales el equilibrio perfecto en subjuego puede existir. Sin embargo, algunas cuestiones importantes quedan abiertas:

1. Las hipótesis hechas en este artículo excluyen que las funciones de transición en juegos sean deterministas. ¿Pueden condiciones más débiles sobre el mecanismo de transición tener en cuenta transiciones deterministas?

2. En este trabajo Mertens y Parthasarathy no muestran que el equilibrio perfecto en subjuegos, por ellos construido, también sea Markoviano, aunque el medio primitivo del juego considerado si sea Markoviano. ¿Puede demostrarse la existencia de un equilibrio Markoviano bajo estas condiciones?

La primera de estas cuestiones es, quizás, la pregunta abierta más importante pendiente en este área, una respuesta afirmativa haría aumentar considerablemente las aplicaciones de los Juegos Markovianos. Esto hace probable que condiciones débiles sobre los mecanismos de transición puedan estar acompañadas de algunas otras condiciones fuertes sobre las funciones de coste (ganancia). Condiciones y resultados más recientes para equilibrios en Juegos Markovianos pueden encontrarse en Balbus and Nowak (2004) y Hernández-Lerma and Guo (2009).

1.2.1 Notaciones y Definiciones

Descripción Formal de un Juego Markoviano (GM)

Un Juego Markoviano se describe por la upla $G = [J, S, (A_i, \Phi_i, r_i, \delta_i)_{i \in N}, q, T]$, donde:

- $J = \{1, 2, ..., n\}$ es el conjunto finito de jugadores.
- S subconjunto Boreliano de algún espacio Polish (espacio métrico completo y separable), es el espacio de los estados del juego.
- Cada jugador $i \in J$ está caracterizado por cuatro elementos $(A_i, \Phi_i, r_i, \delta_i)$, donde:
  - A_i subconjunto Boreliano de algún espacio Polish, es el espacio de acciones del jugador i. Se denotan por a cada uno de los elementos del conjunto de acciones A:
  $$A = \prod_{i \in J} A_i$$
  - Φ_i es una correspondencia de S en A_i, describe para cada estado $s \in S$ el conjunto $\Phi_i(s)$ de todas las posibles acciones que puede tomar el jugador i en el estado s.

- r_i función acotada de $S \times A$ en $\mathbb{R}$, especifica para cada estado s y acción $a \in A$ tomada por el jugador i en s, una recompensa (coste o beneficio) $r_i(s,a)$ para el jugador i.
- $\delta_i \in [0,1]$ es el factor de descuento o de reducción para jugador i.

- q especifica la ley de movimiento o probabilidades de transición para el juego, asociando con cada $(s,a) \in S \times A$ una probabilidad $q(\cdot|s,a)$ sobre el conjunto Boreliano S.
- $T \in \{1, 2, ...\} \cup +\infty$ es el horizonte temporal del juego.

En la formulación más general del Juego Markoviano se pueden considerar acciones aleatorias definidas de la forma siguiente: Sea $M_i(s)$ el conjunto de acciones mixtas disponibles para el jugador i en el estado s, siendo $\Delta(X)$ el conjunto de todas las medidas de probabilidad sobre un conjunto Boreliano X, tenemos que: $M_i(s) = \Delta[\Phi_i(s)]$.

También se define $M(s) = M_1(s) \times \cdots \times M_n(s)$, donde los elementos de $M_i(\cdot)$ y $M(\cdot)$ se denotan por μ_i y μ, respectivamente.

El pago esperado de i por el vector de acciones mixtas μ en el estado s es:

$$r_i(s,\mu) = \int_A r(s,a) d\mu_1(a_1) \cdots d\mu_n(a_n)$$

Para completar la descripción de las componentes, denotamos la probabilidad de transición esperada, por:

$$q_i(\cdot|s,\mu) = \int_A q(\cdot|s,a) d\mu_1(a_1) \cdots d\mu_n(a_n)$$

Informalmente, un Juego Markoviano se desarrolla de la siguiente forma: al principio de cada período de tiempo $t \in [0, 1, \ldots, T-1]$, cada jugador $i \in J$ observa el estado s_t al inicio del período t, y elige una acción $\mu_{it} \in M_i(s_t)$, por convención de notación el primer período de tiempo es el período 0, entonces el último período de un modelo con horizonte finito T es $T-1$. Esta elección se hace con el conocimiento de toda la historia del juego hasta t y de las estrategias de los otros jugadores. Posteriormente a la acción escogida, ocurren dos cosas. La primera, el jugador i recibe la recompensa $r_i(s_t, \mu_t)$, siendo $\mu_t = (\mu_{it})_{i \in N}$ el vector de acciones elegidas. Segundo, el estado transita al valor s_{t+1} correspondiente al siguiente período $t+1$, de acuerdo con la distribución $q(\cdot|s_t, \mu_t)$. La situación se repite nuevamente hasta que se alcanza el período final $T-1$.

El objetivo de cada jugador, en la formulación más clásica, es seleccionar una estrategia que maximice su recompensa total descontada esperada sobre el horizonte del modelo. Un equilibrio resulta cuando ningún jugador puede mejorar unilateralmente con una desviación del vector de estrategias.

Hipótesis Básicas

Para formalizar la estructura de un Juego Markoviano, de modo que sea analíticamente tratable, deben imponerse algunas condiciones de regularidad, que son:

A1 Para todo i, el conjunto A_i es compacto y $\Phi_i : S \to A_i$ es continua en S.

A2 Para todo i, la función r_i es acotada y conjuntamente medible en (s, a) y es continua sobre A para cada estado fijo $s \in S$.

A3 Para cada Boreliano $B \subset S$, la función $q(B|\cdot, \cdot)$ es conjuntamente medible en (s, a) y débilmente continuo sobre A para cada s fijo, es decir, si $a_n \to a$ entonces la sucesión de medidas $q(\cdot|s, a_n)$ converge débilmente a $q(\cdot|s, a)$.

1.2.2 Historias, Estrategias y Equilibrios

Asumimos, excepto donde se indique lo contrario, que el juego tiene un horizonte infinito. Con simples modificaciones, todas las definiciones de esta sección pueden ser extendidas para incluir problemas de horizonte finito.

Historias

Una *t-historia* de un juego es una descripción completa de la evolución del juego hasta el principio del período t. Mediante una t-historia se específica el estado s_τ que ha ocurrido en cada período previo $\tau \in \{0, 1, ..., t-1\}$, las acciones $a_\tau = (a_{i\tau})_{i \in N}$ tomadas por los jugadores en estos períodos y el estado s_t al principio del período t.

Sea H_t el conjunto de todas las posibles t-historias, con h_t un elemento de H_t, entonces h_t es un vector de la forma:

$$h_t = (s_0, a_0, s_1, ..., s_{t-1}, a_{t-1}, s_t).$$

Dado cualquier t y cualquier t-historia, podemos denotar por $s[h_t]$ el estado del t-período s_t que resulta bajo la t-historia h_t.

Estrategias

Una estrategia σ_i para el jugador i es una sucesión $\{\sigma_{it}\}$ donde para cada t y para cada t-historia h_t hasta t, σ_{it} especifica la acción $\sigma_{it}(h_t) \in M_i(s[h_t])$ a ser tomada por el jugador i en el período t como una función medible de la historia h_t hasta t.

Consideramos Σ_i el conjunto de todas las estrategias para el jugador i, y $\Sigma = \times_{i \in N} \Sigma_i$, cada elemento Σ se denota por σ. El vector $(\bar{\sigma}_i; \sigma_{-i})$ es el perfil σ con la estrategia σ_i del i-ésimo jugador reemplazada por $\bar{\sigma}_i$.

Equilibrio

Cada perfil de estrategias σ define en un período t la recompensa esperada $r_{it}(\sigma)(s)$ para el jugador i como función del estado inicial s. De este modo, cada perfil σ define para cada s una recompensa total descontada esperada:

$$W_i(\sigma)(s) = \sum_{t=0}^{\infty} \delta_i^t r_{it}(\sigma)(s) \tag{1.1}$$

Un equilibrio de Nash, o simplemente equilibrio de un juego es un perfil de estrategia σ^* tal que ningún jugador pueda ser beneficiado por una desviación unilateral del perfil, en otras palabras, tal que:

$$W_i(\sigma^*)(s) \geq W_i(\bar{\sigma}_i; \sigma^*_{-i})(s) \qquad \forall s \in S, \quad \bar{\sigma}_i \in \Sigma_i, \quad i \in N \tag{1.2}$$

Equilibrio Perfecto en Subjuego

Sea σ^* un perfil de estrategia de equilibrio. Cada t-historia h_t, no necesariamente consistente con σ^*, induce mediante σ^* un perfil de estrategia para los jugadores para el resto del juego. Denotamos este perfil por $\sigma^*[h_t]$. El equilibrio σ^* se dice que es un *equilibrio perfecto en subjuego* si para todo t y para cada t-historia h_t, la estrategia $\sigma^*[h_t]$ constituye un equilibrio para el resto del juego desde h_t.

La condición de perfección en subjuego es un criterio mínimo de racionalidad para fijar perfiles de estrategia de equilibrio. Un equilibrio que no es perfecto puede recomendar comportamiento de continuación para algunos agentes que es inconsistente con la racionalidad individual.

Estrategias Markovianas

Se observa que en cada período t, la historia pasada del juego continúa influyendo en las posibles recompensas de los jugadores sólo mediante sus efectos sobre el valor del estado s_t en el t-período.

Una *estrategia Markoviana* para el jugador i es una estrategia σ_t tal que para cada t y para cada historia h_t hasta t, σ_{it} dependen de h_t sólo a través del estado en el t-período $s[h_t]$. Como estrategia puede representarse por una sucesión de funciones medibles $\{\pi_i^t\}$ donde, para cada t, π_i^t es una función de S a $\Delta(A_i)$ que satisface que $\pi_i^t \in M_i(s)$ para cada t y para cada $s \in S$.

Si una estrategia Markoviana satisface la condición: $\pi_i^t = \pi_i^\tau (= \pi_i)$, $\forall t, \tau$ entonces se llama *estrategia Markoviana estacionaria* o simplemente estrategia estacionaria. Denotamos por π_i a cada una de las estrategias simples. Finalmente llamamos Π_i^M al conjunto de todas las estrategias Markovianas disponibles para el jugador i, respectivamente Π_i será el conjunto de todas las estrategias Markovianas estacionarias disponibles para el jugador i.

Equilibrio Perfecto de Markov (MPE)

Un equilibrio perfecto en subjuego, en el cual todos los jugadores utilizan sólo estrategias Markovianas es llamado *Equilibrio Perfecto de Markov*, abreviadamente *MPE*.

Cualquier equilibrio en estrategias estacionarias de Markov es necesariamente también perfecto en subjuego. No es necesariamente cierto para equilibrios en estrategias Markovianas no estacionarias. De este modo el requerimiento de perfección en subjuego en la definición de MPE no es vacuo excepto en el contexto de estrategias estacionarias.

1.2.3 Comentarios sobre el entorno

En esta sección se puntualiza que el entorno de Juegos Markovianos descrito es suficientemente general para abarcar como casos particulares muchos de los modelos populares de Economía y Teoría de Juegos, tales como las áreas de Programación Dinámica Estacionaria y Juegos Repetidos.

Programación Dinámica Estacionaria

Estos modelos constituyen el marco matemático más simple y popular en la Teoría Económica Dinámica. Entre la gran variedad de temas que utilizan este marco están: desarrollo económico, equilibrios dinámicos generales, inversión de empresas, tasación de capital, la teoría de investigación y acoplamiento y la adquisición óptima de información. La estructura de Juego Markoviano esbozada constituye una mínima y natural generalización de este marco desde escenarios unipersonales hasta n-personales, y nos permite movernos desde modelos del agente representativo a modelos que permiten una noción más rica de interacción entre agentes.

Por ejemplo, en los modelos de aprendizaje con refuerzo unipersonales, se destaca la naturaleza crítica del equilibrio del agente económico entre maximizar la ganancia actual y la adquisición de información que podría mejorar futuras ganancias posibles. En un escenario multi-agente se puede encontrar el problema *free-rider* que aumenta si todos los agentes pueden aprender algo de cualquier experiencia de algún agente particular. Este es un tema de especial importancia cuando la información no es un bien público pero en su lugar favorece menos al poseedor cuando otros agentes también la poseen.

Juegos Repetidos

Este entorno ha quedado como el modelo preminente para el estudio de interacciones dinámicas estratégicas. Los Juegos Markovianos también constituyen una simple y natural generalización de los Juegos Repetidos. Se puede considerar el Juego Markoviano como una especificación, para cada estado s, de un juego en forma normal $G(s)$ en el cual el pago al jugador i debido a la acción posible $a \in \Phi_1(s) \times \cdots \times \Phi_n(s) \subset A$ viene dado por $r_i(s, a)$. En particular, cuando S es un conjunto unitario $S = \{s\}$, más generalmente, cuando para algún estado s la probabilidad $q(\cdot|s, a)$ es degenerada en s para todo a, el Juego Estocástico para el estado s se reduce a una repetición indefinida del juego $G(s)$ en forma normal con n jugadores que se mueven simultáneamente.

De este modo, los tradicionales Juegos Repetidos con acciones observables forman un caso especial del entorno descrito. No obstante la incorporación de una variable estado en el juego repetido también permite consideraciones de situaciones más generales.

Como casos particulares de Juegos Repetidos se clasifican según el acceso a la información de los jugadores en:

- **Información Imperfecta**. Los jugadores no pueden observar las acciones tomadas por otros agentes en cada período, en su lugar, cada agente observa la realización de una variable aleatoria disponible cuya distribución en cada período se determina parcialmente por el vector de acciones tomadas por los jugadores en el período.

- **Información Perfecta**. Los jugadores se mueven en un orden pre-especificado. Incluidos en esta categoría están los modelos dinámicos de oligopolio en los cuales las firmas se mueven alternativamente, donde hay un conjunto infinito numerable de jugadores $J = \{1, 2, \ldots\}$ y el jugador $i \in J$ se mueve sólo en el período de tiempo t.

Estas clases de juegos pueden ser tratados como un caso particular de Juegos Markovianos si generalizamos las condiciones iniciales, se imponen restricciones sobre el modelo y requiere además que las estrategias de cada jugador dependan sólo de la historia del estado previo y de las acciones previas del jugador.

Juegos Markovianos no Estacionarios

La descripción previa de Juegos Markovianos está centrada en modelos estacionarios con horizonte infinito, pero los modelos no estacionarios, en particular con horizonte finito, son también de nuestro interés, debido a que cada juego siempre se puede convertir en un juego estacionario con horizonte infinito sin pérdida en continuidad o compacidad de la estructura subyacente. Suponemos que el juego está definido por los siguientes elementos:

- S es el espacio de estados.
- T (posiblemente igual a ∞) es su horizonte.
- A_i el espacio de acciones del i-ésimo jugador. Generalmente son conjuntos compactos.
- $\Phi_i^t(s) \subset A_i$ es el conjunto de acciones posibles para el jugador i en el estado s en el período de tiempo t.
- $r_i^t(s, a)$ es la recompensa recibida por el jugador i si las acciones a son tomadas en el estado s en el período de tiempo t. Generalmente son funciones acotadas.
- $\delta_i \in [0, 1)$ el factor de descuento o reducción para el jugador i, es permisible $\delta_i = 1$ si el horizonte temporal T es finito.
- $q^t(\cdot|s, a)$ es la distribución del estado en el $(t + 1)$-período de tiempo si la acción a se toma en el estado s en el período de tiempo t.

Para convertir este juego en un juego estacionario con horizonte infinito, se expande el espacio de estados para incluir el tiempo como una de las coordenadas, el nuevo espacio de estados es $Z = S \times \{1, 2, \ldots, T + 1\}$

Para $\Phi_i(z) = \Phi_i^t(s)$, $z = (s,t) \in Z$ con $a \in A$, además:

$$r_i(z,a) = \begin{cases} r_i^t(s,a) & \text{si} \quad t \neq T+1 \\ 0 & \text{en caso contrario} \end{cases}$$

Definimos las nuevas probabilidades de transición, considerando $(s,t) \in Z$ con $t \in \{1,\dots,T\}$, para el Boreliano $B \subset S$ y $\tau \in \{1,\dots,T+1\}$, esto es:

$$q(B \times \{\tau\}|(s,t),a) = \begin{cases} q^t(B|s,a) & \text{si} \quad \tau = t+1 \\ 0 & \text{si} \quad \tau \neq t+1 \end{cases}$$

Para $t = T+1$, se define $q(B \times \{T+1\}|(s,t),a) = p(B)$ donde $p(\cdot)$ es una medida de probabilidad sobre S.

Este nuevo juego definido representa fielmente al anterior, por construcción, ya que el nuevo juego es un Juego Markoviano estacionario con horizonte infinito y que las propiedades de continuidad del juego original no se pierden con la conversión.

1.2.4 Existencia de Equilibrio Perfecto de Markov

En un entorno Markoviano estacionario, es natural buscar equilibrio en la clase de estrategias estacionarias Markovianas. Los siguientes cinco pasos, como argumento heurístico, se presentan para justificar que el MPE estacionario no es difícil de encontrar.

1. Sea $\pi \in \Pi$ un vector de estrategias estacionarias para los jugadores. En búsqueda de la mejor respuesta para π, un jugador genérico i se enfrenta a un problema de Programación Dinámica Estacionaria con espacio de estados S, espacio de acciones $\Delta(A_i)$, la acción factible correspondiente $M_i(s) = \Delta(\Phi_i(s))$, la función de ganancia

$$r_i(\pi)(s,\mu_i) := r_i(s,(\pi,(s),\mu_i))$$

y las probabilidades de transición:

$$q_i(\pi)(\cdot|s,\mu_i) := q(\cdot|s,(\pi,(s),\mu_i))$$

2. Un recorrido por resultados conocidos de la Teoría de Programación Dinámica nos recuerda que estos problemas poseen una solución si y sólo si admiten una solución Markoviana estacionaria.

3. Por tanto, en busca de la mejor respuesta para $\pi \in \Pi$, el jugador i no pierde flexibilidad estratégica restringiendo su búsqueda a Π_i.

4. Bajo restricciones apropiadas sobre π, el conjunto de las mejores respuestas estacionarias del jugador i para π, llamado $BR_i(\pi)$, se puede garantizar que es no vacío.

5. De este modo podemos llegar a obtener un MPE estacionario del Juego Markoviano como un punto fijo de una aplicación de (algún subconjunto de) Π en sí mismo.

Cuando todos los espacios son finitos se puede obtener un equilibrio en estrategias Markovianas utilizando argumentos de punto fijo. Si se admiten espacios de cardinalidad arbitraria los problemas aumentan de complejidad. Estos argumentos implican que la existencia de MPE en el caso general, sólo puede asegurarse imponiendo algunas estructuras adicionales sobre los problemas originales.

1.2.4.1 Equilibrios Aproximados

Condiciones menos restrictivas que A1-A3 de la Sección 1.2.1 llevan a la existencia de equilibrios aproximados o *quasi-equilibrios*, tales como los ϵ-*equilibrio* y los p-*equilibrio*.

Un ϵ-*equilibrio* es un perfil de estrategias en el cual cada jugador obtiene un pago que está en un entorno de radio ϵ de su mejor respuesta a las estrategias del resto de los jugadores, donde $\epsilon \geq 0$ es un número real fijo. Dicho perfil se puede entender como un equilibrio aproximado, siendo ϵ la medida del grado de aproximación.

Un p-*equilibrio* es un perfil de estrategia en el que cada estrategia de un jugador es la mejor respuesta a las de los demás jugadores, para casi todo estado inicial respecto a la medida p sobre el espacio de estados.

Existen condiciones técnicas en Parthasarathy and Sinha (1989) que permiten obtener MPE a partir de p-equilibrios.

1.3 Desarrollo Sostenible y Juegos Dinámicos

En Hardin (1968) se presenta el clásico juego conocido por la Tragedia de los Comunes, su autor el biólogo Garrett Hardin lo describe con el siguiente escenario, una aldea en la que comparten en común los pastos para alimentar el ganado propiedad de cada familia, es de esperarse que cada pastor racional concluye que la única decisión sensata para él es añadir otro animal a su rebaño, y otro más ..., intentará aumentar ilimitadamente la cantidad de cabezas de ganado para maximizar su ganancia, sin tener en cuenta que el resto de los pastores tendrá el mismo objetivo y que la superficie de pastos es limitada. Como consecuencia del exceso de ganados pastando, en poco tiempo comenzará a escasear el pasto, con lo cuál disminuirá el peso y la cantidad de ganado, siendo la ruina para todos los pastores.

Los problemas de Desarrollo Sostenible aplicados a los recursos naturales de propiedad común, como la contaminación ambiental, la posibilidad de extinción de especies de animales por sobreexplotación, la deforestación de extensas zonas de selva, etc, pueden ser descritos de forma similar a la Tragedia de los Comunes.

Principios Básicos

El Desarrollo Sostenible tiene distintos significados, la definición mas frecuente pertenece al informe Brundtland, también llamado Nuestro Futuro Común, muy relacionada por su significado con la tragedia de los comunes, aparecido en 1987 y se define como:

"El desarrollo sostenible es el desarrollo que satisface las necesidades del presente sin comprometer la capacidad de las generaciones futuras de satisfacer sus propias necesidades".

Su objetivo es mejorar la calidad de vida de todos los ciudadanos de la Tierra, sin aumentar el uso de recursos naturales, más allá de la capacidad del medio ambiente de proporcionarlos indefinidamente. Este concepto busca conciliar el desarrollo económico, el bien social y la protección del medio ambiente. Se trata de tomar acciones, de cambiar políticas y prácticas en todos los niveles, desde el ámbito individual hasta el internacional.

En la Cumbre de la Tierra de Rio en 1992 se desarrolló el marco del informe Brundtland para crear acuerdos y convenciones de problemas críticos como el cambio climático, la desertización y la deforestación. También se hizo el bosquejo de una estrategia amplia de acción (Agenda 21) como el plan de trabajo para los asuntos del medio ambiente y del desarrollo durante las próximas décadas. A lo largo de los 90, se han generado planes de sustentabilidad regionales y sectoriales. Una gran variedad de grupos (desde el sector comercial y gobiernos municipales hasta organizaciones internacionales) han adoptado el concepto y le han dado sus propias interpretaciones particulares. Estas iniciativas han aumentado nuestra comprensión de qué significa el desarrollo sostenible dentro de muchos contextos diferentes.

A nivel esuropeo, el Tratado de Maastricht de 1993 hace referencia al Desarrollo Sostenible en su artículo nº 2: “la comunidad tiene por misión promover un crecimiento durable y no inflacionista respetando el medio ambiente y de inventar modos de desarrollo y de consumo, para asegurar el presente bien común, sin comprometer el bien común del mañana”.

El modelo económico hoy dominante plantea que la economía va bien cuando crece el producto interior bruto (PIB), sin tener en cuenta cuánto cuesta a la colectividad en términos ecológicos y sociales el crecimiento de un punto de PIB, ni que la capacidad de crecimiento económico es finita, ni tampoco tiene en cuenta las limitaciones del sistema natural.

La justificación del desarrollo sostenible proviene tanto del hecho de tener unos recursos naturales limitados, susceptibles de agotarse, como por el hecho de que una creciente actividad económica sin más criterio que el económico actual produce problemas medioambientales tanto a escala local como planetario graves, que pueden en el futuro tornarse irreversibles.

El informe del Programa de las Naciones Unidas para el Medioambiente (PNUMA), en PNUMA (1999), destaca: “el tiempo para una transición racional, bien planificada hacia un sistema sostenible se está acabando rápidamente”(1). Y todavía, continuamos adoptando un enfoque de ”negocios como de costumbre” para tomar decisiones, lo que aumenta la posibilidad de que nuestros sistemas globales se rompan y se derrumben.

Mandatos del Desarrollo Sostenible

Son un llamamiento para cambiar nuestras acciones y hacer las cosas de modo diferente. Cuando estas acciones se toman de forma conjunta, nos pueden ayudar a orientarnos en un camino hacia el desarrollo sostenible. En particular, subrayan la necesidad de:

- Producir de forma diferente: aplicar conceptos de eco-eficiencia y modos de vida sostenibles.
- Consumir de forma diferente: reducir nuestras huellas ecológicas mientras logramos una buena calidad de vida para todos.
- Organizarnos de forma diferente: aumentar la participación pública mientras reducimos la corrupción y subsidios perversos.

Obstáculos principales

- **Complejidad**: El número de factores (jugadores y estrategias) a considerar y la necesidad de tener en cuenta diversas disciplinas (Economía, Ciencias de la Naturaleza, Tecnologías, etc).

- **Incertidumbre**: Los procesos globales no son lineales y dependen del tiempo, lo que implica distintos escenarios y técnicas de estadística y de optimización.

- **Cambio de Paradigma**: Se abandona el paradigma de Entrada/Salida, en favor de la búsqueda de objetivos o la toma de decisiones (racionalidad). La multidimensionalidad de los objetivos y la respuesta adaptativa a los acontecimientos mientras éstos evolucionan en el tiempo, hacen que el estado del sistema de investigación de objetivos sea, en principio, no cognoscible, en cualquier tiempo futuro.

Más información acerca de este tema se puede consultar en von Weizsacker et al. (1997), WBCSD (1999), PNUMA (1999) y en IPCC (2007), entre otros.

1.3.1 Modelos de Desarrollo Sostenible

A partir de los años 70 comienzan a aplicarse técnicas de Teoría de Juegos, Control Óptimo y Programación Dinámica, a modelos de Desarrollo Sostenible, en particular a nuevas formas de consumo, de producción, a la conservación del medio ambiente, etc. A continuación se citan algunos trabajos de diferentes campos, entre muchos que han aparecido en los últimos años, a modo de ejemplo.

- Clemhout and Wan (1985) en *Dynamic common-property resources and environmental problems* y otros trabajos posteriores, como Sorger (2005) y Gaudet and Lohoues (2008), analizan el problema de la sobre-explotación pesquera en ecosistemas marinos incorporando impacto estocástico sobre la interacción natural entre las especies, utilizan técnicas de juegos diferenciales, obtienen un equilibrio del nivel de pesca de forma proporcional a la reproducción de las especies, para no agotar el stock de recursos. Entre los miles de trabajos relacionados con la conservación de las especies marinas, queremos mencionar algunos más recientes. En Martín-Herrán and Rincón-Zapatero (2005) se aplican las condiciones necesarias y suficientes para caracterizar la eficiencia del equilibrio de Nash por medio de un sistema de ecuaciones diferenciales parciales cuasilineales. En Bailey et al. (2010) encontramos una revisión de los trabajos realizados en los últimos años sobre la aplicación de la Teoría de Juegos a la explotación pesquera de forma racional.

- Murphy et al. (1989) en *A Dynamic Nash Game Model of Oil Market Disruption and Strategic Stockpiling*, presentan y analizan un modelo de juego dinámico para analizar políticas publicas y privadas para el manejo de inventarios y tarifas en mercados con probabilidad de fallo en el suministro y reservas de petroleo en Estados Unidos, donde cada jugador escoge su política individulamente para lograr maximizar el beneficio esperado, incluyen condiciones para la unicidad y la estabilidad del equilibrio. También presentan resultados relacionados con la determinación de un subjuego perfecto para determinar el equilibrio en un modelo con horizonte temporal infinito. Como continuación de este modelo encontramos el trabajo de Wu et al. (2008) aplicado a un país importador de petroleo como es China, asumiendo incertidumbre en el precio del barril lo cual influye directamente sobre la cantidad a adquirir en el mercado mundial.

- Germain et al. (1996), en *Calcul économique itératif et stratégique pour les negociations internationales sur les pluies acides entre la Finlande, la Russie et l'Estonie*. Estudian el enfoque del problema de la lluvia ácida como forma cooperativa internacional, utilizan un modelo no lineal y procesos de aprendizaje, en un juego repetido con horizonte infinito, tienen en cuenta las emisiones por períodos, pero no el stock de contaminación. En este trabajo pretenden lograr un óptimo cooperativo utilizando transferencias financieras entre los países implicados con el objetivo de reducir los daños causados al medioambiente, este trabajo representa un antecedente directo de Germain et al. (2003) and Eyckmans and Tulkens (2003), ambos muy comentados en esta memoria.

- Kanudia and Shukla (1998) en *Modelling of Uncertainties and Price Elastic Demands in Energy-Environment Planning for India*, describen dos variantes del modelo técnico económico MARKAL, implementados en el caso de India. MARKAL es un modelo técnico-económico orientado al estudio integrado de sistemas energéticos que describe el proceso de oferta de energía, endogenizando las demandas internas. La primera variante es el uso de la programación estocástica para incluir futuras incertidumbres en el análisis. La segunda variante es la inclusion de funciones de demanda sensibles al precio. En el análisis incorpora futuras incertidumbres para estudiar la reducción de las emisiones de carbón y determina el efecto en los precios derivados de diferentes políticas con respecto al control ambiental, la formulación del modelo la realizan en forma lineal en un período corto de tiempo. Un modelo parecido al anterior aplicado a

Canadá, aparece en Kanudia and Loulou (1997), donde obtienen los costes marginales, los costes totales, etc.

- Gabriel et al. (2003) en *Computational Experience with a Large-Scale, Multi-Period, Spatial Equilibrium Model of the North American Natural Gas System* y Gabriel et al. (2005) en *A large-scale linear complementarity model of the North American natural gas market* describen el Gas Systems Analysis Model (GSAM) desarrollado por ICF Consulting para el U.S. Department of Energy (DOE) para analizar el mercado del gas norteamericano. GSAM es un modelo de gran escala que calcula el equilibrio del mercado máximizando el excedente del productor menos el coste del transporte del gas, con 12 regiones como productores y 14 como demandantes. GSAM se ha utilizado extensivamente en análisis del sector estadounidense de gas natural para contestar preguntas en lo que concierne a políticas para el abastecimiento, la demanda, y el transporte de gas.

- La optimización de recursos hidraúlicos y su conservación se analizan en Madani (2009) como un juego dinámico con dos jugadores en tres períodos de tiempo. En Dechert and O'Donnell (2006) resuelven numéricamente el problema de contaminación de las aguas de un lago, planteado como un problema de juegos dinámicos estocásticos con una cantidad N de comunidades o jugadores y horizonte temporal infinito, cada estado del sistema representa el nivel acumulado de contaminantes cuya dinámica de transición depende del estado actual y de una variable aleatoria multiplicativa sobre la variable de control, en función de la cantidad de lluvia caída en cada período de tiempo.

La cantidad de trabajos que utilizan técnica de juegos, control o programación dinámica, estocásticos aplicados a problemas de Desarrollo Sostenible ha crecido exponencialmente en los últimos años, a pesar de la complejidad de estas herramientas, pero estos modelos describen dichos problemas de una forma más cercana.

1.4 Objetivos y Estructura de la Tesis

Los problemas de optimización dinámica que se plantean, resuelven, implementan e ilustran con un ejemplo basado en un modelo económico medioambiental real (RICE), en esta tesis, responden a un doble objetivo. Por una parte, capturar los elementos de incertidumbre presentes en el problema de control del stock de contaminación, mejorando las modelizaciones utilizadas hasta ahora, lo cual se puede considerar como el objetivo teórico. Existe además un segundo objetivo de índole práctica.

Los paradigmas tradicionales de comportamiento de los países frente a las posibles negociaciones en temas de contaminación ambiental son: el cooperativo, donde todos los países de forma voluntaria tratan de resolver un único problema (ver modelo $P1$ de la Sección 2.3) y el no cooperativo (ver modelo $P2$ de la Sección 2.4). Las sucesivas conferencias internacionales sobre contaminación ambiental y cambio climático han echado por tierra la verosimilitud de la cooperación voluntaria, por lo que la siguiente preocupación en escenarios de negociación fué la búsqueda de compensaciones que eventualmente incentivaran la cooperación. Una de ellas es sin duda la compensación monetaria, la cual abrió las puertas a los mercados de emisiones negociables, no habiendo producido aún todos los equilibrios esperados.

De acuerdo con esta nueva estrategia, en esta tesis se analiza un problema de control del nivel acumulado (stock) de contaminación ambiental, que formula la posibilidad de transferencias monetarias, etapa a etapa para cada país con el fin de incentivar la cooperación entre los países (ver modelo $P3$ de la Sección 3.2.3).

Por último y como alternativa unilateral al equilibrio de Nash que se obtiene como solución no cooperativa del problema $P2$ de la Sección 2.4, se presenta para cada país una posible estrategia que permita globalmente rebajar el nivel acumulado (stock) de contaminación ambiental, ver modelo $P4$ de la Sección 4.2. La estrategia óptima de cada país se obtiene como una solución de un problema de maximización con función objetivo probabilística que involucra un valor objetivo (target) que es fijado por cada país dentro del rango de los costes marginales correspondientes a las soluciones cooperativas de modelo $P1$ y no cooperativas del modelo $P2$. De este modo la solución óptima obtenida, se puede contemplar, desde un punto de vista práctico, como un elemento a ofrecer al país que se supone inmerso en un proceso de negociación internacional con el objetivo de disminuir la contaminación ambiental y frenar el cambio climático.

En la idea de poder comparar las distintas soluciones óptimas y funciones de valor correspondientes que se obtienen como soluciones de los problemas $P1$-$P4$, se ha considerado interesante definir un único escenario de aplicación, que permitiera ilustrar las características de los óptimos obtenidos. Para ello se ha tomado como referencia el propuesto por Eyckmans and Tulkens (2003) como escenario común a todas las secciones de ilustración numérica contenidas en esta memoria, ver Secciones 2.5, 3.4 y 4.4.2.

El resto de la memoria se estructura como sigue. En el Capítulo 2 se presenta el modelo de Control del nivel acumulado de contaminación ambiental con sus componentes. Se modelizan las dos formas de comportamiento de los agentes (países o regiones) implicados, el problema cooperativo y el problema no cooperativo, se estudian sus soluciones y para cada uno se desarrolla un código que permite la aproximación numérica a la solución. Para seis países o regiones durante un período de tiempo de 40 años, se presentan los resultados numéricos obtenidos de emisiones de CO_2, del stock y de los costes en que incurren cada uno con las soluciones obtenidas, haciendo una comparación ambos resultados. Además se hace un análisis de diferentes tipos particulares de la función de daños.

En el Capítulo 3, se define el concepto de transferencias con el objetivo de motivar a todos los países a participar en la gran coalición para obtener el óptimo cooperativo, se presenta un nuevo modelo no cooperativo, se estudia el planteamiento y la solución del problema de optimización correspondiente considerando las transferencias obtenidas a partir de las soluciones resultantes de los modelos cooperativo $P1$ y el no cooperativo $P2$ del Capítulo 2.

En el Capítulo 4 se estudia el modelo de control con criterios de probabilidad en la función objetivo, de forma tal que se minimiza la probabilidad de que el coste marginal de cada país no exceda de un valor (target) fijado previamente, se enuncian las propiedades relativas a este tipo de problema y la existencia de soluciones del mismo. Se obtienen resultados numéricos para el problema simulado. Utilizando el modelo descrito en el Capítulo 2 se genera un código con las características de este nuevo problema y se ilustran los resultados con el ejemplo ya mensionado.

Por último, en el Capítulo 5, se resumen las principales aportaciones de la tesis y se comentan las futuras líneas de investigación.

Capítulo 2

A Stochastic Model for the Environmental Transnational Pollution Control Model

"¿Cómo osamos hablar de leyes del azar? ¿No es, acaso, el azar la antítesis de cualquier ley?"

Bertrand Rusell

Abstract

In this chapter we provide a stochastic dynamic game formulation of the transnational pollution control problem when the environmental damage arises from stock pollutant that accumulates, for accumulating pollutants such as CO_2 in the atmosphere. With a few exceptions the dynamic games in the literature facing this problem operates in a deterministic framework. Although the objective difficulty of modelling complex interactions among environmental and meteorological factors influencing the dynamic of the stock pollutant, we propose a stochastic dynamic model where the inherent uncertainty of the cumulated stock pollutant's evolution due to those factors is considered. Thus, we calculate the optimal path of abatement as the solution of the stochastic game for cooperative and non-cooperative behavior of the countries. The optimality criteria assumed in our setting is the minimization of the expected discounted total cost. To illustrate our proposal following the RICE model we

present some numerical results based on real scenarios for six regions, USA, Japan, European Union, China, Former Soviet Union and the Rest of the World.

2.1 Introduction

In the last years, atmospheric pollution has been of great concern for many countries of the world, since the exposure to air pollution is associated with numerous effects on human health. One way to tackled this issue focus on the analysis and modelization of the general framework of the individual's exposure, see for instance Zidek et al. (2007) and references therein.

Due to the international dimension of the global environmental resources, the transboundary effects of the polluting activities and the prospect of climate change, a second approach focused on the analysis of the requirements for an optimal solution based on game theory has been developed. It is the strong correlation between wealth production and maintenance of wealthy lifestyles on the one hand, and energy consumption and consequently pollution on the other hand, that makes them the historical responsible for the current environmental situation.At the heart of the matter lie economic considerations. How can the supra-national regulator then best tread the frontier between wealth producing activities and the carrying of controlling the international stock pollutant?

The literature has initially analyzed the problem in a cooperative perspective, and the theory on international environmental agreements (IEA) and the prospect of climate change has motivated many game theoretical studies, focused on cooperation and core solutions, and seek for the mechanisms, mostly transfers or exchanges, able to sustain cooperation between countries. The necessity of cooperation amongst the countries involved, if a social optimum is to be achieved, has already been addressed in the literature in terms of Game Theory concepts; see e.g. Barrett (2003), Finus (2001) and references therein for a review on these topics. With a few exceptions this literature works with simple static models of pollution despite the fact that many of the important environmental problems, as climate change, the depletion of the ozone layer or the acid rain problem, are caused by a stock pollutant. However, the stock of pollution may change in the course of the game, as a result of a positive rate of natural decay and emissions of the countries. Thus, the presence of a stock pollutant leads to a dynamic game that is not strictly repeated. The stability of an International Environmental Agreement among n countries that emit pollutant are studied using differential games, defined in continuous time, by Jorgensen et al. (2003) and Jorgensen et al. (2004), Rubio and Casino (2005), among others.

Due to the absence of a supra-national authority able to boost the emergence and implementation of a coordinated international pollution abatement agreements, a second stream focused on the formation of environmental agreements from the non-cooperative game theory, has appeared in the literature. Although there is a substantial literature that uses game-theoretic and optimal control concepts to analyze stock pollutant control, there are only a few attempts modeling that issue in a stochastic framework. Contributions that stand out in this stream are Carraro and Filar (1995), Barrett (2003), Haurie and Viguier (2003), Eyckmans and Tulkens (2003), Germain et al. (2003), and Fuentes-Albero and Rubio (2010) for an interesting overview.

An appropriate modeling is a crucial issue to evaluate stock pollutant and to prepare plans and programs requested by the framework EU Directive on air quality assessment and management (96/62/EC) and related directives. The choice of the approach to model the dynamic of the stock pollutant variable should consider besides human actions interactions among several environmental factors. As Fontanella et al. (2007) pointed out, the variety of pollutants may undergo chemical reactions themselves and with other species, and the pollutants move as a result of transport by the wind and they diffuse as a result of turbulence in the air. Furthermore, the diffusion of pollutant is often a result of two interconnected factors such as meteorological conditions and local topography. All these issues show how complex the stock pollutant phenomenon is and how a deterministic scenario may result in a poor modeling of its evolution.

The use of stochastic control models to develop climate-economy models has been advocated by Haurie and Viguier (2003) to represent the possible competition between Russia and China on the international market of carbon emissions permits, their model includes a representation of the uncertainty concerning the date of entry of developing countries on this market in the form of an event tree. Also by Bahn et al. (2008), they show how a piecewise deterministic stochastic control model, over an infinite time horizon, can be used as a paradigm for the design of efficient climate policy, their model recognizes the existing uncertainty concerning the true sensitivity of climate, and the fact that the solution to the climate change issue may reside in the introduction of new carbon-free technologies. Keller et al. (2004) have already explored the combined effects of uncertainty and learning about a climate threshold (an uncertain ocean thermohaline circulation collapse) in an economic optimal growth model.

In this chapter, a stochastic theoretical framework for the stock pollutant control is constructed in a similar setting as the deterministic one in Germain et al. (2003). A discrete-time model with a finite planning period (horizon) for n countries is proposed and solve under the cooperative and the non-cooperative paradigms. We consider stock of pollutant in a wide sense, not restricted to the carbon dioxide (CO_2) stock level. Inclusion of manifold pollutants is important. To wit, the 1997 Kyoto Protocol to the Framework Convention on Climate Change limits aggregate emissions of six direct greenhouse gases, such as: carbon dioxide (CO_2), methane (CH_4), nitrous oxide (N_2O), hydrofluorocarbons ($HFCs$), perfluorocarbons ($PFCs$), sulphur hexafluoride (SF_6)), as well for the indirect greenhouse gases such as SO_2, NOx, CO and/or micro particles of industrial pollution (between 0.1 y 2.5 μ-meters). The emissions are aggregated and considered as CO_2 equivalents. We consider that the only way to control the stock of pollution is through the control of emissions, that is reducing pollution is done through the reduction of emissions, and not through the cleaning of the environment. In addition, we assume that emissions are proportional to production, see Jorgensen and Zaccour (2001) among many others.

As far as we know, none stochastic formulation for the finite horizon dynamic analysis of international agreements on transnational pollution control has been introduce as an extension of the issues presented in Germain et al. (2003). We adopt this point of view because to consider randomness on the factors in the model is closer to reality (see Casas and Romera (2009a)).

With our methodology, the cooperative and non-cooperative models are formulated as Markov Decision Processes (MDP) with constraints, for an overview of the topic we refer to the reader to Hernández-Lerma (1999) and Puterman (2005). By applying stochastic dynamic programming techniques, i.e. see Bertsekas (2000), then we solve the corresponding optimization problems and obtain the optimal emission strategies, the optimal cumulated stock pollutant and the optimal value functions, under cooperative and non-cooperative scenarios. The criteria to be minimized is the expected total discounted cost which consider two issues, the damages caused by the stock pollutant to each country's environment measured in monetary terms, and the cost for the country to reduce its own emissions.

This chapter is organized as follows. In Section 2.2 the international stock pollutant model formulation, the modes of countries behave, and the description of the underlaying MDP formulation are introduced. In Section 2.3, we describe the international stock pollutant control cooperative model and we solve the optimization problem which provides the optimal trajectories of emissions and stock which constitute the *international optimum.* In Section 2.4 we solve the n optimization problems obtained when countries do not sign a voluntary IEA. In Section 2.5, a numerical example based on real scenarios borrowed from Eyckmans and Tulkens (2003) are presented and regarding the solutions some comments on the use of these optima are included. Section 2.6 summarizes conclusions and extensions of this chapter.

2.2 Stock Pollutant Control Model

In this section, we formulate a discrete-time stochastic dynamic game with finite planning horizon. We refer to the reader to the modelization in Germain et al. (2003) for a deterministic counterpart of our setting.

We consider a Markovian Game described by a tuple

$$G = \{J, \mathcal{T}, S, E, p, \}, \tag{2.1}$$

with the following elements; n players where $J = \{1, 2, ..., n\}$ denotes the set of countries (regions). A finite planing horizon with discrete-time periods t, such that $t \in \mathcal{T} = \{1, 2, ..., T\} \subset \mathbb{Z}^+$. The state space of the game, S, with element s, is a Borel subset of some Polish (i.e., complete, separable, metric) countable and non empty space. The control variables or actions (emissions) are $e_{it} \in E$, where E is the countable and non empty overall control space or action space, and $E = \bigcup_{s \in S} E(s)$, where $E(s)$ is the finite set of *admissible actions* (emissions), when the system is in each state (pollutant level) $s \in S$. The law of motion (or transition probabilities) p for the game defined for each $(s, e) \in S \times E$ is the conditional probability $p(\cdot|s, e)$ over the Borel sets of S.

The state of the system is the accumulated level of pollution in the atmosphere given as stock of pollutant at each period t, $s_t \in S$, which evolves according to the state equation

$$s_t = (1 - \delta)s_{t-1} + \sum_{i=1}^{n} e_{it} + \xi_t \quad , \qquad 1 \leq t \leq T, \tag{2.2}$$

where s_0 is the initial stock of pollutant or preindustrial level, given, and δ is the pollutant's natural rate of atmospheric absorption of CO_2 between two periods of time, such that $0 < \delta <$

1. The random disturbance ξ_t is a noise process: a sequence of independent and identically distributed (i.i.d.) random variables, and independent of the initial state s_0, with

$$\mathbb{E}\left[\xi_t\right]=0, \qquad \sigma^2=\mathbb{E}\left[\xi_t^2\right]<\infty, \qquad \forall t=1,2,...,T-1. \tag{2.3}$$

We consider that future costs are discounted by the constant and positive *discount factor* β with $0<\beta\leq 1$. Function $c_i(e_{it})$ measures in monetary terms the total cost incurred by country $i\in J$ at period $t\in\mathcal{T}$ from limiting its own industrial emissions to e_{it}; is a differentiable, decreasing ($c_i'<0$) and strictly convex function ($c_i''>0$). Function $d_i(s_t)$ measures in monetary terms the damages caused by the stock of pollutant s_t during the time period t for the i-th country; is a differentiable, increasing ($d_i'>0$) and convex function ($d_i''\geq 0$). The damages in each country's environment depend on the emissions of pollutant of all different countries at each time-period t that contribute to a stock s_t.

In cooperative form the countries jointly choose at each period its emissions levels in order to minimize the expected total discounted costs, then the resulting trajectories of emissions and stock constitute the *international optimum.* In non-cooperative form, each country considers only the damages of the stock of pollutant over itself. In the sense of a Nash equilibrium, the countries minimize, at each period, only its own expected total discounted cost, with knowledge of the emissions vector e_{jt}, with $j\neq i$, of the other countries.This is the standard setting based on average costs optimality.

The decision problems considered in this work can be formulated as a discrete-time, finite-horizon and stationary MDP with expected total discounted cost. Then, we can express the elements of our random scenarios through the following MDP

$$\Gamma=(S,E,R,P,\beta), \tag{2.4}$$

where the *state space* S and the overall *action space*

$$E=\bigcup_{s\in S}E(s)$$

have been previously defined. The *cost set* R is a bounded countable subset of $\mathbb{R}$. For each $t\geq 1$, let s_t, e_t and r_t denote the state (pollutant level) of the system, the action (emissions) taken by the decision maker (pays), and the cost incurred at period of time t, respectively. The stationary, single-stage, conditional *transition probabilities* are defined by

$$p^e_{i,j,r}:=Prob\left(s_{t+1}=j,r_t=r/s_t=i,e_t=e\right),$$
$$\forall i,j\in S,\quad r\in R,\quad t\geq 1,\sum_{j\in S,r\in R}p^e_{i,j,r}=1,\quad i\in S,\quad e\in E(i).$$

2.3 Cooperative Model

One assumes that the countries behave in an internationally optimal way. We solve the following problem

$$
\begin{aligned}
(P1) \qquad & \min_{\{e_{it}\}} \quad && \mathbb{E}\left[\sum_{t=1}^{T}\sum_{i=1}^{n}\beta^t\left[c_i(e_{it})+d_i(s_t)\right]\right] \\
& \text{s.t.} && s_t=(1-\delta)s_{t-1}+\sum_{i=1}^{n}e_{it}+\xi_t \\
& && e_{it}\geq 0, \qquad \forall i=1,\cdots,n; \quad \forall t=1,\cdots,T \\
& && s_0>0
\end{aligned}
$$

The resulting family of trajectories of emissions e_{it}^W for all players $i\in J$ determined together with the resulting stock s_t^W, constitute the international optimum for all periods $t\in\mathcal{T}$ or a cooperative equilibrium, see for instance Dutta and Sundaram (1998).

Note that the objective function in $(P1)$ is equivalent to

$$
\min_{\{e_{it}\}}\mathbb{E}\left[\sum_{t=1}^{T}\sum_{i=1}^{n}\beta^t\left[c_i(e_{it})+d_i(s_t)\right]\right]\Leftrightarrow\min_{\{e_{it}\}}\sum_{t=1}^{T}\sum_{i=1}^{n}\beta^t\left(c_i(e_{it})+\mathbb{E}\left[d_i(s_t)\right]\right) \tag{2.5}
$$

Proposition 2.1. *Problem* $(P1)$ *has an equilibrium* $\{e_{it}^*\}$.

Proof. The convexity of the functions $c_i(e_{it})$ and $d_i(s_t)$, for all $i\in J$ and for all periods $t\in\mathcal{T}$, suffices to guarantee that the minimum exists and is unique, see for instance, Puterman (2005) or Hernández-Lerma (1999). □

In a multi-criteria optimization context the former problem $(P1)$ has been solved by Hernández-Lerma and Romera (2004).

This problem $(P1)$ can be solved by using Stochastic Dynamic Programming tools. The *expected value function* W, according to Bellman's principle of optimality, satisfies the Dynamic Programming equations for $(P1)$ subject to (2.2)

$$(P1.1) \qquad W(T, s_{T-1}) = \min_{e_{iT}} \mathbb{E}\left[\sum_{i=1}^{n} (c_i(e_{iT}) + d_i(s_T))\right],$$

$$(P1.2) \qquad W(t, s_{t-1}) = \min_{e_{it}} \mathbb{E}\left[\sum_{i=1}^{n} [c_i(e_{it}) + d_i(s_t)] + \beta W(t+1, s_t)\right].$$

$$\forall t = 1, 2, ..., T-1$$

$$\begin{array}{lrcl} \text{s.t.} & s_t & = & (1-\delta)s_{t-1} + \sum_{i=1}^{n} e_{it} + \xi_t \\ & e_{it} & \geq & 0 \qquad \forall t \in \mathcal{T} \quad , \quad \forall i \in J \\ & s_0 & > & 0 \end{array}$$

The stochastic Dynamic Programming equations $(P1.1)$ and $(P1.2)$ are equivalent, respectively, to

$$(P1.1) \;\; \Leftrightarrow \;\; W(T, s_{T-1}) = \min_{e_{iT}} \left\{ \sum_{i=1}^{n} (c_i(e_{iT}) + \mathbb{E}\left[d_i(s_T)\right]) \right\}$$

$$(P1.2) \;\; \Leftrightarrow \;\; W(t, s_{t-1}) = \min_{e_{it}} \left\{ \sum_{i=1}^{n} (c_i(e_{it}) + \mathbb{E}\left[d_i(s_t)\right]) + \beta W(t+1, s_t) \right\}$$

If countries cooperate, they jointly solve $(P1.1)$ at final period of time T, the country i's expected total cost is

$$W_i(T, s) = c_i(e_{iT}^W) + d_i(s_T^W), \qquad \text{with} \qquad s_T^W = [1-\delta]s + \sum_{i=1}^{n} e_{iT}^W$$

where $(e_{1T}^W, e_{2T}^W, \ldots, e_{nT}^W)$ is the vector of optimal emission levels or policy, and s_T^W denotes the resulting stock of pollutant at final period T, and s is the inherited stock of pollutant at the begin of period T.

In earlier periods, if countries cooperate they solve the problem $(P1.2)$ for $1 \leq t \leq T-1$. Optimal levels of emissions and resulting stock of pollutant are denoted by e_{it}^W and s_t^W respectively. Then let denotes the country i's expected total discounted equilibrium cost by

$$W_i(t, s) = c_i(e_{it}^W) + d_i(s_t^W) + \beta W_i(t+1, s_t^W), \qquad \text{with} \qquad s_t^W = [1-\delta]s + \sum_{i=1}^{n} e_{it}^W,$$

where s is the inherited stock of pollutant at the begin of period t.

Let define as τ-expected discounted total cost by

$$W_i^\tau \equiv \sum_{t=1}^{\tau} W_i(t, s_{t-1}^W), \quad 1 \leq \tau \leq T-1,$$

and the total cost

$$W_i \equiv \sum_{t=1}^{T} W_i(t, s_{t-1}^W). \tag{2.6}$$

Note that (2.6) is the marginal cost incurred by country i under the cooperative paradigm.

2.3.1 Cooperative Alternative Problem

We present an equivalent cooperative problem which can be solved by using Linear Programming tools. The recurrence equation (2.2), of the stock pollutant s_t, gives a dynamic character to the cooperative model. We may write an associated model, by writing s_t as a function of the known initial stock s_0, the emissions e_{it} from each country $i \in J$, and the random disturbance vector ξ_t in each period of time $t \in \mathcal{T}$.

From (2.2) we obtain the following recursive expression for the stock pollutant s_t

$$\begin{aligned}
s_1 &= (1-\delta)s_0 + \sum_{i=1}^{n} e_{i1} + \xi_1. \\
s_2 &= (1-\delta)^2 s_0 + (1-\delta)\sum_{i=1}^{n} e_{i1} + (1-\delta)\xi_1 + \sum_{i=1}^{n} e_{i2} + \xi_2. \\
&\cdots \\
s_t &= (1-\delta)^t s_0 + (1-\delta)^{t-1}\sum_{i=1}^{n} e_{i1} + (1-\delta)^{t-1}\xi_1 + \cdots + \\
&\quad + \cdots + (1-\delta)^{t-\tau}\sum_{i=1}^{n} e_{i\tau} + (1-\delta)^{t-\tau}\xi_\tau + \cdots + \sum_{i=1}^{n} e_{it} + \xi_t.
\end{aligned}$$

Thus, the general form is

$$s_t = (1-\delta)^t s_0 + \sum_{\tau=1}^{t}(1-\delta)^{t-\tau}\sum_{i=1}^{n} e_{i\tau} + \sum_{\tau=1}^{t}(1-\delta)^{t-\tau}\xi_\tau. \tag{2.7}$$

Explicitly developing the previous recurrence equation, we obtain the following system of restrictions for the optimization problem $(P1)$,

$$
\begin{aligned}
s_1 &= (1-\delta)s_0 + e_{11} + e_{21} + \cdots + e_{n1} + \xi_1. \\
s_2 &= (1-\delta)^2 s_0 + (1-\delta)e_{11} + (1-\delta)e_{21} + \cdots + \\
&\qquad +(1-\delta)e_{n1} + (1-\delta)\xi_1 + e_{12} + \cdots + e_{n2} + \xi_2. \\
&\cdots \\
s_t &= (1-\delta)^t s_0 + (1-\delta)^{t-1}e_{11} + \cdots + (1-\delta)^{t-1}e_{n1} + (1-\delta)^{t-1}\xi_1 + \cdots + \\
&+(1-\delta)^{t-2}e_{12} + \cdots + (1-\delta)^{t-2}e_{n2} + (1-\delta)^{t-2}\xi_2 + \cdots + (1-\delta)e_{1t-1} \\
&+ \cdots + (1-\delta)e_{nt-1} + (1-\delta)\xi_{t-1} + e_{1t} + \cdots + e_{nt} + \xi_t.
\end{aligned}
$$

By using the **Markov's condition or the Property of causality**, $\forall j, r \in \{0, 1, \ldots, N-1\}$ with $j < r$, it is shown that the state x_r only depends on the state x_j and the intermediate controls $\{u_j, u_{j+1}, \ldots, u_{r-1}\}$. Then, we conclude that the actual contamination stock depends on the initial stock s_0 and the set of controls or emission vector $e_1, e_2, ..., e_T$ for each period of time $t \in \mathcal{T}$.

Note that, by definition, $e_{it} \geq 0$ and $s_t \geq 0$, for all $t \in \mathcal{T}$ provided that $0 < \delta < 1$. Then, we can consider equivalently the following problem with convex objective function and $T+1$ linear constraints

$$
\begin{aligned}
\min_{\{e_{it}\}} \quad & \mathbb{E}\left[\sum_{t=1}^{T}\sum_{i=1}^{n}\beta^t\left[c_i(e_{it}) + \tilde{d}_i(s_0, e_{it}, \xi_t)\right]\right] \\
\text{s.t.} \quad Ae &= b - \xi \\
e &\geq 0 \\
s_0 &> 0
\end{aligned}
$$

$$
\begin{aligned}
\text{where } e' &= (e_{11}; e_{21}; \cdots; e_{n1}; e_{12}; e_{22}; \cdots; e_{n2}; \cdots; e_{1T}; e_{2T}; \cdots; e_{nT}) \\
b' &= \left(-(1-\delta)s_0; -(1-\delta)^2 s_0; \cdots; -(1-\delta)^T s_0\right) \\
\xi' &= \left(\xi_1; (1-\delta)\xi_1 + \xi_2; \ldots; \ldots; (1-\delta)^{T-1}\xi_1 + (1-\delta)^{T-2}\xi_2 + \cdots + \xi_T\right)
\end{aligned}
$$

The independent vector b and random disturbance vector ξ are of order T. The matrix

A is a $T \times Tn$, lower triangular matrix, with the following structure

$$A = \begin{pmatrix} 1 & \cdots & 1 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ (1-\delta) & \cdots & (1-\delta) & 1 & \cdots & 1 & \cdots & 0 & \cdots & 0 \\ (1-\delta)^2 & \cdots & (1-\delta)^2 & (1-\delta) & \cdots & (1-\delta) & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ (1-\delta)^{T-1} & \cdots & (1-\delta)^{T-1} & (1-\delta)^{T-2} & \cdots & (1-\delta)^{T-2} & \cdots & 1 & \cdots & 1 \end{pmatrix}$$

By using the development presented in this section one can find the solutions e_{it}^W of optimal emissions for each country $i \in J$, and one obtains the stock levels of contamination s_t^W in each period of time $t = 1, 2, ..., T$.

2.3.2 Analysis of particular Damage Functions

Note that although the cost function c_i, depends only on the emissions e_{it} of each country $i \in J$ at each period of time $t \in \mathcal{T}$, the damages function d_i depends on the initial stock s_0, the emissions of the each one others countries e_{it} with $i \neq j$, the emissions e_{it} and the random disturbance ξ_t, for each period of time t. This fact determines the stochastic structure of the objective function to be considered in the optimization problem $(P1)$, as it is shown in (2.5).

We analyze useful cases of damage functions that appear in the economic literature, and we present the particular programming problems to be solved in each case. This analysis remains valid for both models, cooperative and non cooperative with some slight modification.

2.3.2.1 Linear Case

One assume

$$d_i(s_t) = as_t + b, \qquad a, b \in \mathbb{R}.$$

Following (2.5) the objective function of the cooperative model $(P1)$ has the following form

$$\min_{\{e_{it}\}} \quad \left\{ \sum_{t=1}^{T} \sum_{i=1}^{n} \beta^t \left(c_i(e_{it}) + a\mathbb{E}\left[d_i(s_t) \right] \right) \right\}$$

because

$$\begin{aligned}
\mathbb{E}\left[d_i(s_t)\right] &= \mathbb{E}\left[as_t + b\right], \\
&= a\mathbb{E}\left[s_t\right] + b, \\
&= a\mathbb{E}\left[(1-\delta)^t s_0 + \sum_{\tau=1}^{t}(1-\delta)^{t-\tau}\sum_{i=1}^{n} e_{i\tau} + \sum_{\tau=1}^{t}(1-\delta)^{t-\tau}\xi_\tau\right] + b, \\
&= a\left[(1-\delta)^t s_0 + \sum_{\tau=1}^{t}(1-\delta)^{t-\tau}\sum_{i=1}^{n} e_{i\tau} + \mathbb{E}[\sum_{\tau=1}^{t}(1-\delta)^{t-\tau}\xi_\tau]\right] + b, \\
&= a\left[(1-\delta)^t s_0 + \sum_{\tau=1}^{t}(1-\delta)^{t-\tau}\sum_{i=1}^{n} e_{i\tau}\right] + b.
\end{aligned}$$

Then the objective function of model $(P1)$ is equal to the objective function of the following linear programming

$$\min_{\{e_{it}\}} \quad \left\{ \sum_{t=1}^{T} \sum_{i=1}^{n} \beta^t \left(c_i(e_{it}) + a(1-\delta)^t s_0 + a\sum_{\tau=1}^{t}(1-\delta)^{t-\tau}\sum_{i=1}^{n} e_{i\tau} + b \right) \right\}.$$

2.3.2.2 Quadratic Case

One assume that

$$d_i(s_t) = (s_t)^2.$$

The objective function of the cooperative model has the following form

$$\min_{\{e_{it}\}} \quad \left\{ \sum_{t=1}^{T} \sum_{i=1}^{n} \beta^t \left(c_i(e_{it}) + \mathbb{E}\left[(s_t)^2\right] \right) \right\}.$$

then

$$\begin{aligned}
\mathbb{E}\left[(s_t)^2\right] &= \mathbb{E}\left[\left((1-\delta)^t s_0 + \sum_{\tau=1}^{t}(1-\delta)^{t-\tau}\sum_{i=1}^{n} e_{i\tau} + \sum_{\tau=1}^{t}(1-\delta)^{t-\tau}\xi_\tau\right)^2\right], \\
&= \mathbb{E}\left[\varphi^2 + \left(\sum_{\tau=1}^{t}(1-\delta)^{t-\tau}\xi_\tau\right)^2 + 2\varphi\left(\sum_{\tau=1}^{t}(1-\delta)^{t-\tau}\xi_\tau\right)\right], \\
&= \varphi^2 + \mathbb{E}\left[\left(\sum_{\tau=1}^{t}(1-\delta)^{t-\tau}\xi_\tau\right)^2\right] + 2\varphi\mathbb{E}\left[\sum_{\tau=1}^{t}(1-\delta)^{t-\tau}\xi_\tau\right].
\end{aligned}$$

where

$$\varphi = (1-\delta)^t s_0 + \sum_{\tau=1}^{t}(1-\delta)^{t-\tau}\sum_{i=1}^{n} e_{i\tau}.$$

Provided that $\{\xi_t\}$ are i.i.d. and condition (2.3), we have

$$\begin{aligned}
\mathbb{E}\left[\left(\sum_{\tau=1}^{t}(1-\delta)^{t-\tau}\xi_\tau\right)^2\right] &= \mathbb{E}\left[\sum_{\tau=1}^{t}(1-\delta)^{2(t-\tau)}\xi_\tau^2 + 2\sum_{\tau=1}^{t}(1-\delta)^{t-\tau}\xi_\tau\sum_{j=1}^{t}(1-\delta)^{t-j}\xi_j\right], \\
&= \sum_{\tau=1}^{t}(1-\delta)^{2(t-\tau)}\mathbb{E}\left[\xi_\tau^2\right] + 2\sum_{\tau=1}^{t}\sum_{j=1}^{t}(1-\delta)^{t-\tau}(1-\delta)^{t-j}\mathbb{E}\left[\xi_\tau\xi_j\right], \\
&= \sum_{\tau=1}^{t}(1-\delta)^{2(t-\tau)}\sigma_t^2, \\
\text{then} \quad \mathbb{E}\left[(s_t)^2\right] &= \varphi^2 + \sum_{\tau=1}^{t}(1-\delta)^{2(t-\tau)}\sigma_t^2.
\end{aligned}$$

Then in this case, our problem is transformed in an quadratic programming problem.

2.3.2.3 Exponential Case

Finally, one assume

$$d_i(s_t) = \exp(s_t).$$

The objective function of the cooperative model has the following form

$$\min_{\{e_{it}\}} \quad \left\{\sum_{t=1}^{T}\sum_{i=1}^{n}\beta^t\left(c_i(e_{it}) + \mathbb{E}\left[\exp(s_t)\right]\right)\right\}.$$

Now

$$\begin{aligned}\mathbb{E}\left[\exp(s_t)\right] &= \mathbb{E}\left[\exp\left((1-\delta)^t s_0\right)\exp\left(\sum_{\tau=1}^{t}(1-\delta)^{t-\tau}\sum_{i=1}^{n}e_{i\tau}\right)\exp\left(\sum_{\tau=1}^{t}(1-\delta)^{t-\tau}\xi_\tau\right)\right]\\ &= \exp\left((1-\delta)^t s_0\right)\exp\left(\sum_{\tau=1}^{t}(1-\delta)^{t-\tau}\sum_{i=1}^{n}e_{i\tau}\right)\mathbb{E}\left[\exp\left(\sum_{\tau=1}^{t}(1-\delta)^{t-\tau}\xi_\tau\right)\right]\end{aligned}$$

where

$$\begin{aligned}\mathbb{E}\left[\exp\left(\sum_{\tau=1}^{t}(1-\delta)^{t-\tau}\xi_\tau\right)\right] &= \prod_{\tau=1}^{t}\mathbb{E}\left[\exp(1-\delta)^{t-\tau}\xi_\tau\right],\\ &= \prod_{\tau=1}^{t}\varphi_\xi\left[(1-\delta)^{t-\tau}\right].\end{aligned}$$

We recognize φ_ξ as the z-transformed function if ξ follows a discrete random variable.

Depending on the expression of this φ_ξ function, we get different types of objective functions, and therefore different types of mathematic programming problems, usually they will be non-linear optimization problems.

2.4 Non-Cooperative Model

In an alternative mode of behaviour, we describe what would happen if the countries do not sign a voluntary international environmental agreement. One may assume that countries behave non cooperatively in the sense of Nash equilibrium, where each of them minimizes at each period only its own discounted costs, taking given the emissions of the other countries. A Nash equilibrium is a family of strategies, one for each player, that minimize every country i's cost, given the strategies of all other players $j \neq i$. In such an equilibrium, no individual country has an incentive to deviate as long as the other countries stick to their equilibrium strategies, see for instance Carraro and Filar (1995) or Barrett (2003).

Formally, there are n problems to solve. Actually, at each period of time $t \in \mathcal{T}$, for each country $i \in J$ we solve the following problem

$$
\begin{aligned}
(P2) \qquad & \min_{\{e_{i\tau}\}} \quad && \mathbb{E}\left[\sum_{\tau=t}^{T} \beta^{\tau}\left[c_i(e_{i\tau}) + d_i(s_{\tau})\right]\right] \\
& \text{s.t.} && s_t = (1-\delta)s_{t-1} + \sum_{i=1}^{n} e_{it} + \xi_t \\
& && e_{it} \geq 0 \quad \forall t \in \mathcal{T}; \quad \forall i \in J \\
& && s_0 > 0
\end{aligned}
$$

Note that the objective function in $(P2)$ is equivalent to

$$
\min_{\{e_{i\tau}\}} \mathbb{E}\left[\sum_{\tau=t}^{T} \beta^{\tau}\left[c_i(e_{i\tau}) + d_i(s_{\tau})\right]\right] \Leftrightarrow \min_{\{e_{i\tau}\}} \sum_{\tau=t}^{T} \beta^{\tau}\left(c_i(e_{i\tau}) + \mathbb{E}\left[d_i(s_{\tau})\right]\right)
$$

Proposition 2.2. *Problem $(P2)$ has an equilibrium $\{e_{it}^N\}$.*

Proof. A particular case the convexity of the functions $c_i(e_{it})$ and $d_i(st)$, for all $i \in J$ and for all periods $t \in \mathcal{T}$, suffices to guarantee that the Nash equilibrium exists and is unique (see for instance, Puterman (2005) and Hernández-Lerma (1999)). □

The *expected value functions* N_i, according to Bellman's principle of optimality, can be found by solving the Stochastic Dynamic Programming equations for $(P2)$ subject to (2.2)

$$
\begin{aligned}
(P2.1) \qquad & N_i(T, s_{T-1}) = \min_{e_{iT}} \mathbb{E}\left[c_i(e_{iT}) + d_i(s_T)\right] \\
(P2.2) \qquad & N_i(t, s_{t-1}) = \min_{e_{it}} \mathbb{E}\left[c_i(e_{it}) + d_i(s_t) + \beta N_i(t+1, s_t)\right]
\end{aligned}
$$

The resulting family of trajectories of emissions (policies) e_{it}^N determined for each country $i \in J$, together with the resulting stock s_t^N, constitute a non-cooperative Nash equilibrium for all periods $t \in \mathcal{T}$ (see Dutta and Sundaram (1998)).

The Stochastic Dynamic Programming equations $(P2.1)$ and $(P2.2)$ are equivalent, respectively, to

$$
\begin{aligned}
(P2.1) \quad \Leftrightarrow \quad & N_i(T, s_{T-1}) = \min_{e_{iT}} c_i(e_{iT}) + \mathbb{E}\left[d_i(s_T)\right], \\
(P2.2) \quad \Leftrightarrow \quad & N_i(t, s_{t-1}) = \min_{e_{it}} c_i(e_{it}) + \mathbb{E}\left[d_i(s_t)\right] + \beta N_i(t+1, s_t).
\end{aligned}
$$

In the non cooperative equilibrium the country i's expected total cost at period final T is

$$N_i(T,s) = c_i(e_{iT}^N) + d_i(s_T^N) \quad \text{with} \quad s_T^N = [1-\delta]s + \sum_{i=1}^{n} e_{iT}^N.$$

where $(e_{1T}^N, e_{2T}^N, \ldots, e_{nT}^N)$ is the vector of emissions equilibrium level, and s_T^N the resulting stock of pollutant at final period T, where s is the inherited stock of pollutant at the begin of period T.

Let define as τ-expected discounted total cost by

$$N_i^\tau \equiv \sum_{t=1}^{\tau} N_i(t, s_{t-1}^N), \qquad 1 \leq \tau \leq T-1,$$

and the total cost

$$N_i^\tau \equiv \sum_{t=1}^{\tau} N_i(t, s_{t-1}^N), \quad 1 \leq \tau \leq T-1 \quad \text{with} \quad N_i \equiv \sum_{t=1}^{T} N_i(t, s_{t-1}^N). \tag{2.8}$$

Note that (2.8) is the marginal cost incurred by country i under the non-cooperative paradigm.

2.4.1 Non-Cooperative Alternative Problem

Following the same technique as in subsection (2.4.1) and bearing in mind the explicit recursive expression (2.7) obtained for the stock pollutant s_t, we have to solve for each country $i \in \{1, 2, ..., n\}$ the alternative problem

$$\begin{array}{rrcl} & \min\limits_{\{e_t\}} & & \mathbb{E}\left[\sum_{t=1}^{T} \beta^t \left[c_i(e_t) + \tilde{d}_i(s_0, e_t, \xi_t)\right]\right] \\ \text{s.a.} & B_i e & = & b_i + \xi \\ & e & \geq & 0 \qquad \forall t \in \mathcal{T} \\ & s_0 & > & 0 \end{array}$$

where

$$\begin{aligned} e_i' &= (e_{i1}; e_{i2}; \cdots ; e_{iT}) \\ b_i' &= (b_{i1}; b_{i2}; \cdots ; b_{iT}) \\ b_{it} &= -(1-\delta)^t s_0 - \sum_{\tau=1}^{t} \sum_{j \neq i}^{n} (1-\delta)^{t-\tau} e_{j\tau} \\ \xi' &= \big(\xi_1; (1-\delta)\xi_1 + \xi_2; \ldots; \ldots; (1-\delta)^{T-1}\xi_1 + (1-\delta)^{T-2}\xi_2 + \cdots + \xi_T\big) \end{aligned}$$

The matrix B_i is a square matrix, lower triangular, of order T, with ones in the principal diagonal. The vector b and the random disturbance ξ have order T. The structure of the matrix B_i is as follows

$$B_i = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \cdots & 0 \\ (1-\delta) & 1 & 0 & 0 & 0 & \cdots & 0 \\ (1-\delta)^2 & (1-\delta) & 1 & 0 & 0 & \cdots & 0 \\ (1-\delta)^3 & (1-\delta)^2 & (1-\delta) & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & 0 \\ (1-\delta)^{T-1} & (1-\delta)^{T-2} & (1-\delta)^{T-3} & \cdots & \cdots & (1-\delta) & 1 \end{pmatrix}$$

then we can may obtain the inverse matrix of the matrix B_i, which is quasi diagonal

$$B_i^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \cdots & 0 \\ -(1-\delta) & 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & -(1-\delta) & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & -(1-\delta) & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & \cdots & -(1-\delta) & 1 \end{pmatrix}$$

As in the cooperative model solution, by using the development presented in this section one can find the path e_{it}^N of optimal emissions for each country $i \in J$, and the optimal stock level of contamination s_t^N at each period of time $t = 1, 2, ..., T$.

Note that the particular analysis for linear, quadratic and exponential damage functions developed in section 3.2.1, holds for the non cooperative case with little change in the objective function.

2.5 A Numerical Example

In the following, we show some numerical results obtained by application of the algorithms developed in the preceding sections of cooperative ($P1$) and non cooperative ($P2$) problems to a real scenario considering six regions or countries. The model and the values of the parameters used are based on the paper by Eyckmans and Tulkens (2003). In that paper the model named the Climate Negotiation (CLIMNEG) World Simulation Model, is considered as well as a deterministic dynamic analysis about how many countries will be interesting in signing an international environmental agreement (IEA) with accumulating pollutant in discrete time. All computations were made by use of the software Matlab 7.3.0 (R2006b).

2.5.1 Model and parameters

The temperature change equation is taken from the climate economy model RICE (Regional Integrated model of Climate and the Economy) by Nordhaus and Yang (1996) and Nordhaus and Boyer (2000), as well as most of the parameter values and all basic data on Gross Domestic Product (GDP), population, capital stock, carbon emissions and concentration and global mean temperature. The division of the world is the same as in the RICE model. There are 6 countries or regions: USA, Japan, European Union (EU), China, Former Soviet Union (FSU) and Rest of the World (ROW). The time is divided in years, the initial period (period $t = 0$) refers to year 1990, following the ONU (1997). To take account on the long term impacts of stock pollutant, we take a long planning horizon of 100 years, but we will only consider results until 2030 in order to avoid boundary problems.

The CO_2 emissions in each region or country $i \in J$ at period of time $t \in \mathcal{T}$ are denoted by e_{it}, with $e_{it} \geq 0$ for all $i \in J$ and for all $t \in \mathcal{T}$, and $e_t = (e_{1t}; e_{2t}; ...; e_{nt})$ is the corresponding vector of emissions of CO_2 in each of n regions or countries i at period of time t. Emissions of region i at time t are considered due to economic activity and proportional to the potential GDP named Y_{it}, according to expression

$$e_{it} = \sigma_{it}(1 - \eta_{it})Y_{it} \tag{2.9}$$

The optimal abatement rate of control of emissions, in each country or region i and in every period of time t, is the endogenous vector $\eta_t = (\eta_{1t}; \eta_{2t}; ...; \eta_{nt})$ with $0 \leq \eta_{it} \leq 1$, for all $i \in J$ and for all $t \in \mathcal{T}$. Note that $\eta_t = 0$ for all t determines the "business-as-usual" (BAU) scenario in this model, i.e. a trajectory in which the emissions are not reduced with respect to their maximum values.

The emissions of CO_2 to output ratio σ_{it}, of each country or region i at each period of time t, declines exogenously over time t due to an assumed autonomous energy efficiency increase. Given e_{it} and Y_{it}, and the BAU scenario, one may obtains

$$\sigma_{i,t} = \frac{e_{it}}{Y_{it}}.$$

The potential GDP denoted by Y_{it} is the output(exogenous) of country or region i at period of time t, in billion 1990 USA dollars, and g_{it} is the annual growth rates of each country or region i at each period of time t.

$$Y_{i,t+1} = (1 + g_{it})Y_{it}. \tag{2.10}$$

The next equation modelizes the stock pollutant part of the model. The emissions contribute to the stock of CO_2 in the atmosphere, in billion tons of carbon CO_2, according to equation (2.2)

$$s_t = (1-\delta)s_{t-1} + \sum_{i=1}^{n} e_{it} + \xi_t, \qquad \forall t = 1, ...T.$$

or equivalently

$$s_t = (1-\delta)^t s_0 + \sum_{\tau=1}^{t}(1-\delta)^{t-\tau}\sum_{i=1}^{n} e_{i\tau} + \sum_{\tau=1}^{t}(1-\delta)^{t-\tau}\xi_\tau.$$

where the initial stock or preindustrial level of the CO_2 atmospheric stock, is taken as 590 billion tons of carbon equivalent.

The parameter δ, such that $0 < \delta < 1$, the rate of decay or absorption of CO_2 in the atmosphere between two periods of time t and $t-1$, is assumed as $\delta = 0.0833$ per decade or $\delta = 0.0909512$ per year.

The random disturbance ξ_t is a noise process as in (2.3), with $\mathbb{E}[\xi_t] = 0$ and $\sigma^2 = \mathbb{E}[\xi_t^2] = 1$, $\forall t = 1, 2, ..., T-1$. In our simulations we averaged the damages function over 100 runs carried out after the corresponding 100 values of the standard normal disturbance ξ_t generated.

The stock s influences in turn the variation of atmospheric temperature w.r.t. its preindustrial or initial level s_0, according to the following equations

$$\Delta T_t = \gamma \ln \left(\frac{s_t}{s_0} \right),$$

where the annual discount rate γ is an exogenous positive parameter. This parameter is calibrated such that a doubling of CO_2 atmospheric concentration results in an increase of temperature of 2.5 degrees with respect to its preindustrial level, and we take its value as

$$\gamma = \frac{2.5}{\ln(2)}.$$

The next two equations describe the economic part of the model, i.e. the costs c_{it} of reducing the emissions of CO_2 on the one hand, and the costs of the damages d_{it} due to stock pollutant and climate change on the other.

The abatement cost function c_{it} of country i at each period of time t, measured in billion 1990 USA dollars, is given by

$$c_{it}(e_{it}) = a_{i1} \eta_{it}^{a_{i2}} Y_{it} = a_{i1} \left[1 - \frac{e_{it}}{\sigma_{it} Y_{it}} \right]^{a_{i2}} Y_{it},$$

where the functions c_{it} are decreasing ($c'_{it} < 0$) and strictly convex ($c''_{it} > 0$), as is assumed in Section 2.2.

Damages due to stock pollutant and climate change are assumed to follow from the increase of the atmospheric temperature, in billion 1990 USA dollars, according to

$$d_{it}(s_t) = b_{i1} \Delta T_t^{b_{i2}} Y_{it} = b_{i1} \left[\gamma \ln \left(\frac{s_t}{s_0} \right) \right]^{b_{i2}} Y_{it}, \tag{2.11}$$

where the functions d_{it} are increasing ($d'_{it} > 0$) and convex ($d''_{it} > 0$), according to the hypotheses of the model in Section 2.2.

The regional exogenous parameter values a_{i1}, a_{i2}, b_{i1} and b_{i2} for all countries $i \in J$ are exogenous and positive, and are given in Table 2.1.

We now describe the exogenous parameters appearing in the problems $(P1)$ and $(P2)$. The initial output Y_{it}, i.e. 1990 potential GDP, of the different region or countries are given by the vector

$$Y_{1990} = [5464.796, 2932.055, 6828.042, 370.024, 855.207, 4628.621],$$

Table 2.1: Regional exogenous parameter values per country

i	USA	JAP	EU	CHI	FSU	ROW
a_{i1}	0.07	0.05	0.05	0.15	0.15	0.1
a_{i2}	2.887	2.887	2.887	2.887	2.887	2.887
b_{i1}	0.01102	0.01174	0.01174	0.015523	0.00857	0.02093
b_{i2}	2.0	2.0	2.0	2.0	2.0	2.0

expressed in billion 1990 USA dollars and the total of the world, at this year, is 21078.750 billions USA dollars.

The average annual output growth rates g_{it} in per cent for each country at each period of time t, given in Table 2.2, are calculated from Kverndokk (1994). After (2.10) it is possible to evaluate Y_{it} for all $i \in J$ and for all $t \in \mathcal{T}$, the cumulative output of region or country i during the period of time t.

Table 2.2: Average annual output growth rates g_{it} in %, per country for each period of time t (per decade)

period t	USA	JAP	EU	CHI	FSU	ROW
1990-2000	2.60	2.20	2.20	4.60	2.60	3.70
2000-2020	2.20	1.70	1.70	4.40	2.10	3.40
2020-2050	1.60	1.30	1.30	3.40	1.60	2.70
2050-2080	1.00	1.00	1.00	2.50	1.00	1.50
2080-2110	1.00	1.00	1.00	2.00	1.00	1.00

We face now the calculation of the initial value σ_{1990} for the optimization problem. The initial CO_2 vector of emissions e_{1990}, in absence of any control are taken from the RICE model and these emissions are measured in billion tons of carbon,

$$e_{1990} = [1.37, 0.29, 0.872, 0.805, 1.066, 3.43].$$

Given e_{1990} and the annual GDP Y_{1990} value, following (2.9) we obtain the initial emissions of CO_2 to output ratio σ_{1990}

$$\sigma_{1990} = [0.0002506, 0.0000989, 0.0001277, 0.0021755, 0.0012464, 0.000741].$$

Given e_{1990} and the annual emissions growth rates g_{it}, following (2.10) it is easy to calculate the output ratio σ_{it} for all country $i \in J$ and for all period of time $t \in \mathcal{T}$, that is the CO_2 emission ratio of region or country i during the period t.

In this example we borrow the output Y_{it} and CO_2 emission/output ratio time series from different versions of the RICE model, developed by Nordhaus and Yang (1996) and Nordhaus and Boyer (2000).

Finally the discount factor per year, that appears in the objective functions of problems $(P1)$ and $(P2)$, where the annual discount rate, with $\rho = 0.02$, is taken as

$$\beta = \frac{1}{(1+\rho)} = 0.98.$$

2.5.2 Numerical Results

In this subsection we present the reference scenario which corresponds to the values of the parameters given in the last subsection. The simulations are made for a time horizon of 100 years, but we give the results only up to 2030, i.e. for the first 40 years, in order to avoid boundary problems.

We have implemented the equivalent formulation of problems $(P1)$ and $(P2)$ given in Sections 2.3.1 and 2.4.1, respectively. The damages function (2.11) considered in our example is more complex, thus, we have developed specific Matlab code for its implementation.

Note that the optimal abatement rates for each country can be directly obtained after the optimal emissions by applying (2.9). This is in fact one of the outputs more frequently analyzed by the economic literature concerning stock pollutant control.

In the sequel, Tables 2.3 and 2.4 and Figures 2.1 and 2.2 are related with cooperative scenario'results from Section 2.3. Tables 2.5 and 2.6 and Figures 2.3 and 2.4 are related with non-cooperative scenario'results from Section 2.4.

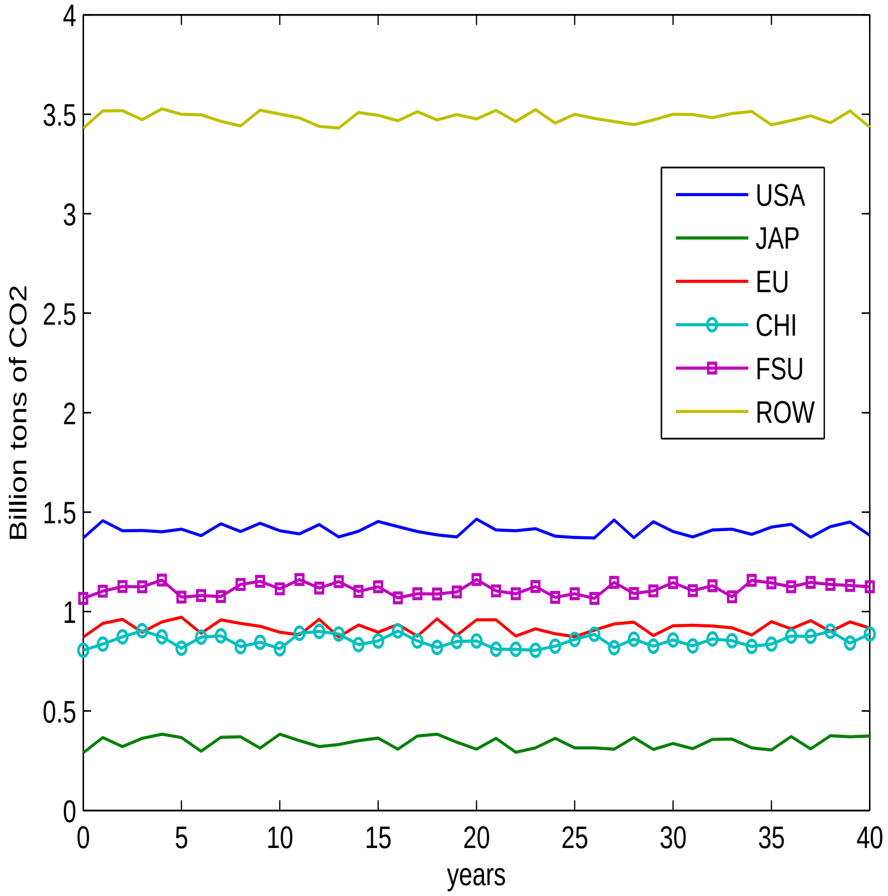

Figure 2.1: Optimal cooperative emissions e_{it}^W for each country at each period of time t in billion tons of carbon equivalent.

Table 2.3 gives the optimal cooperative emissions e_{it}^W in billion tons of CO_2 equivalent for each country during each period of time t. These results are related with problem $(P1)$. The last row gives the cumulated emissions per country until the end of the horizon T in billion tons of carbon. Figure 2.1 shows the optimal cooperative emissions e_{it}^W for each country i and per each period of time t.

Table 2.4 gives the optimal cooperative value function W_{it} for each country during each period of time t in billions of 1990 USA dollars. These results are related with problem $(P1)$. The last row gives the cumulated value function per country and the total of the world at the end of the final period T, measured in billions of 1990 USA dollars. Figure 2.2 shows the optimal cooperative value function W_{it} for each country i and per each period of time t in billions of 1990 USA dollars.

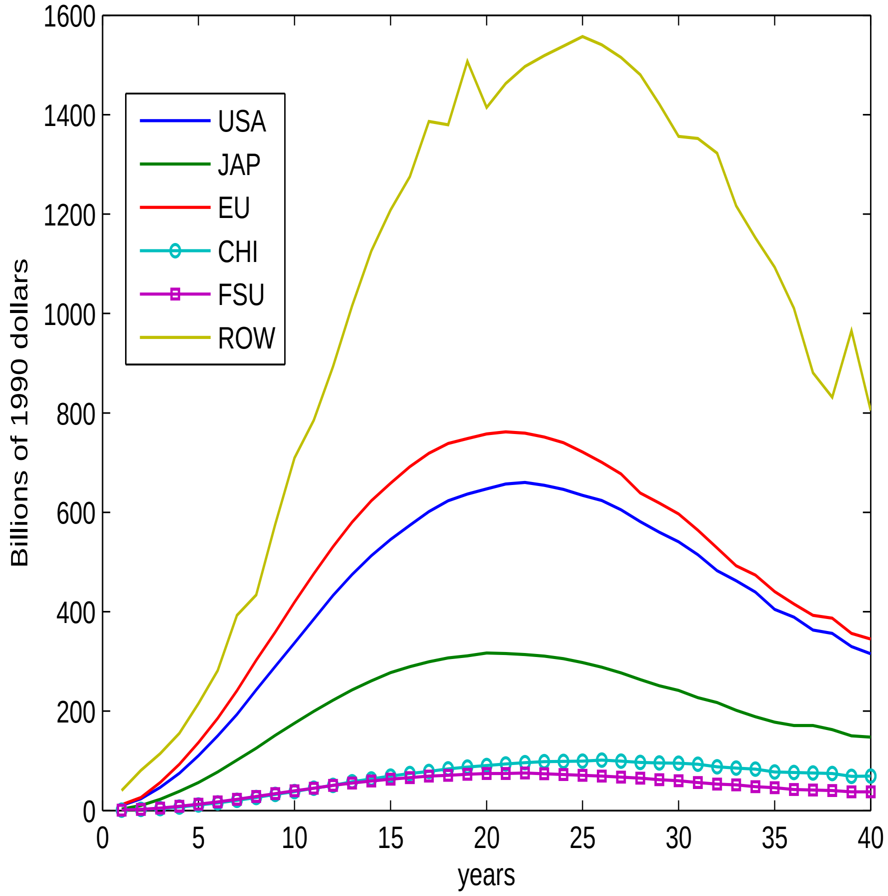

Figure 2.2: Optimal Cooperative Value Function W_{it} per country i for each period of time t in billions of 1990 USA dollars.

Table 2.5 gives the optimal non cooperative emissions e_{it}^N per country i during the period of time t. These results are related with problem $(P2)$. The last row gives the cumulated emissions per country i until the end of the period of time T in billion tons of carbon. Figure 2.3 shows the optimal non cooperative emissions e_{it}^N for each country i and for each period of time t.

Table 2.6 gives the optimal non cooperative value function N_{it} for each country during each period of time t in billions of 1990 USA dollars. These results are related with problem $(P2)$. The last row gives the cumulated value function for each country and the total of the world until the end of the horizon T, measured in billions of 1990 USA dollars. Note that the Figure 2.4 shows the optimal non cooperative value function N_{it} for each country i and per each period of time t in billions of 1990 USA dollars.

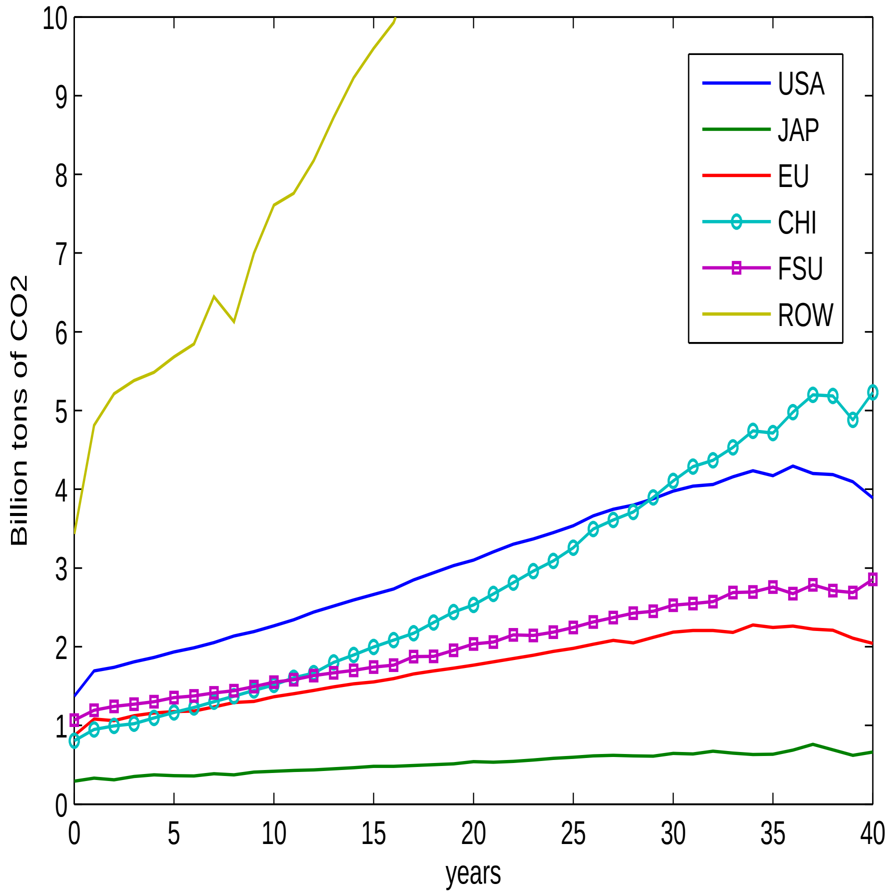

Figure 2.3: Optimal non cooperative emissions e_{it}^N per each country i at each period of time t.

Remark 2.1. *Although optimal emissions increase with time for both cases, in the cooperative case, see Figure 2.1, it is not a remarkable issue. Nevertheless, we discover an increasing trend of the optimal emissions in the non cooperative case, as is shown in Figure 2.3. As it is expected, the total optimal non cooperative emissions for each country given by the last row in Table 2.5 are bigger that the corresponding total optimal cooperative emissions given by the last row in Table 2.3.*

The total Optimal Cooperative Value Function is smaller than the total Optimal Non Cooperative Value Function, as it is shown in Tables 2.4 and 2.6, and Figures 2.2 and 2.4. We observe that this result is consistent with what is obtained in the seminal paper for the deterministic model provided by Germain et al. (2003). In fact this result was expected after the definition of the optimum.

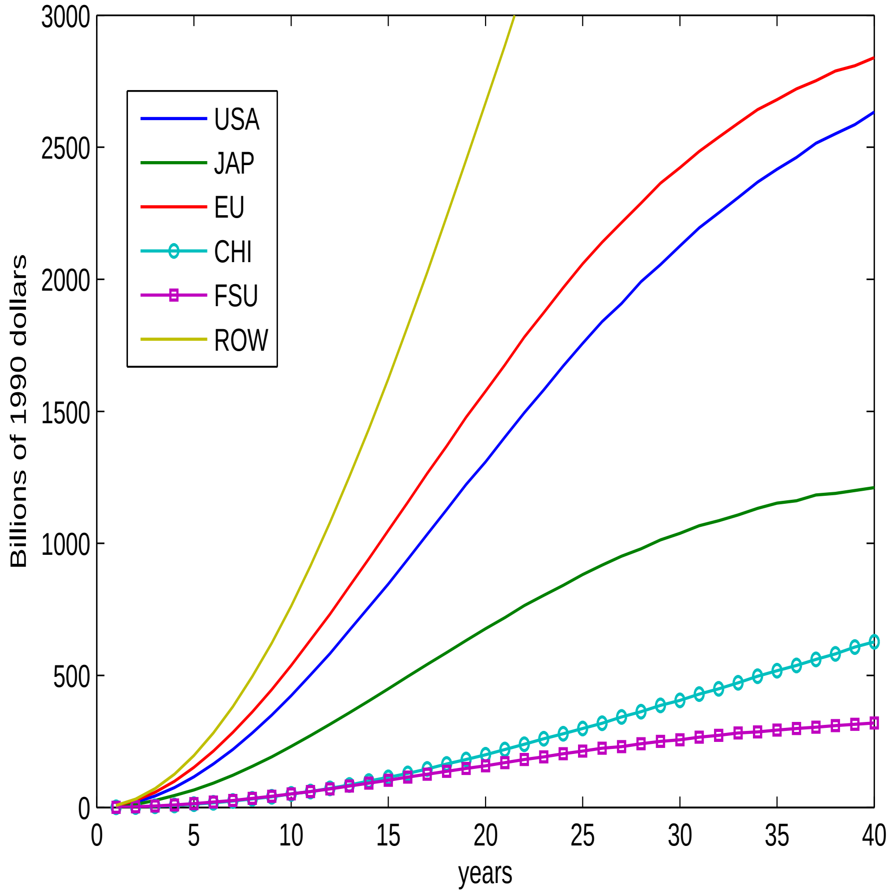

Figure 2.4: Optimal Non Cooperative Value Function N_{it} per country i at each period of time t.

We compare now the optimal stocks of pollutant. Table 2.7 gives the cooperative optimal stock of pollutant, s_t^W, the non cooperative optimal stock of pollutant s_t^N and the differences between them at each period of time t in billion tons of carbon. We observe a great improvement of the cooperative behavior with respect to the non cooperative one over the time.

Figure 2.5 depicts the optimal cooperative and non-cooperative stocks of pollutant, s_t^W and s_t^N respectively for each period of time t in billion tons of carbon equivalent. Note that the optimal stock cooperative s_t^W decreases faster than the non-cooperative stocks s_t^N. This result is consistent with the expected behavior of the solutions of problems $(P1)$ and $(P2)$.

We have checked our model in different scenarios by changing the values of the noise process parameters including the deterministic case, i.e. $E[\xi_t] = 0$, $Var[\xi_t] = 0$. All the

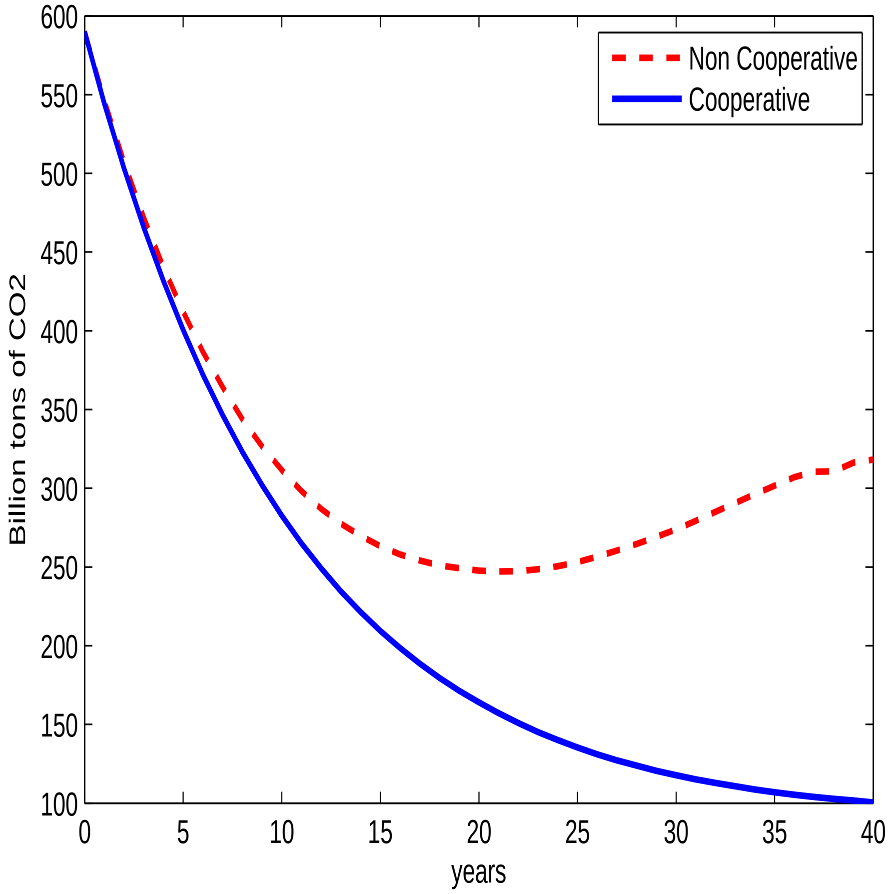

Figure 2.5: Optimal cooperative s_t^W and non-cooperative s_t^N stocks

results we have found were consistent, and for the deterministic case we have obtained optimal stationary strategies for both problems $P1$ and $P2$, as we expected.

Table 2.3: Optimal cooperative emissions e_{it}^W for each country i at each period of time t in billion tons of carbon equivalent.

t	USA	Japan	EU	China	FSU	ROW	Total
0	1.370	0.292	0.872	0.805	1.066	3.430	7.835
1	1.457	0.366	0.940	0.836	1.102	3.516	8.219
2	1.405	0.321	0.961	0.873	1.125	3.518	8.206
3	1.407	0.362	0.896	0.904	1.124	3.472	8.167
4	1.400	0.383	0.948	0.873	1.158	3.526	8.290
5	1.414	0.366	0.971	0.816	1.072	3.500	8.140
6	1.381	0.298	0.891	0.870	1.080	3.496	8.018
7	1.441	0.367	0.957	0.878	1.075	3.464	8.185
8	1.401	0.370	0.939	0.822	1.136	3.441	8.111
9	1.443	0.313	0.926	0.845	1.152	3.520	8.201
10	1.406	0.384	0.896	0.812	1.114	3.501	8.115
11	1.390	0.350	0.883	0.891	1.160	3.480	8.157
12	1.437	0.320	0.960	0.900	1.117	3.438	8.174
13	1.375	0.332	0.872	0.887	1.150	3.431	8.048
14	1.404	0.351	0.932	0.834	1.101	3.508	8.131
15	1.453	0.364	0.896	0.851	1.124	3.494	8.183
16	1.426	0.307	0.934	0.902	1.069	3.466	8.107
17	1.401	0.374	0.876	0.852	1.089	3.512	8.107
18	1.385	0.383	0.963	0.820	1.087	3.470	8.111
19	1.374	0.343	0.880	0.849	1.098	3.498	8.044
20	1.465	0.308	0.958	0.852	1.160	3.476	8.221
21	1.409	0.362	0.959	0.810	1.103	3.519	8.164
22	1.405	0.292	0.877	0.811	1.090	3.463	7.940
23	1.416	0.315	0.912	0.805	1.125	3.522	8.098
24	1.379	0.362	0.888	0.825	1.072	3.455	7.983
25	1.372	0.315	0.874	0.859	1.089	3.499	8.011
26	1.370	0.315	0.906	0.886	1.066	3.478	8.023
27	1.460	0.308	0.937	0.818	1.146	3.463	8.135
28	1.370	0.366	0.946	0.860	1.090	3.447	8.082
29	1.451	0.307	0.879	0.826	1.103	3.471	8.039
30	1.402	0.337	0.928	0.857	1.145	3.499	8.170
31	1.375	0.311	0.931	0.828	1.105	3.498	8.049
32	1.409	0.358	0.927	0.862	1.130	3.482	8.170
33	1.413	0.358	0.918	0.852	1.073	3.503	8.120
34	1.387	0.314	0.881	0.824	1.157	3.513	8.079
35	1.423	0.304	0.948	0.835	1.144	3.446	8.103
36	1.438	0.371	0.912	0.876	1.124	3.468	8.192
37	1.374	0.309	0.954	0.874	1.146	3.491	8.151
38	1.426	0.375	0.899	0.900	1.136	3.456	8.196
39	1.450	0.371	0.947	0.842	1.131	3.516	8.259
40	1.382	0.374	0.917	0.887	1.124	3.434	8.121
Total	56.397	13.701	36.838	34.121	44.608	139.371	325.039

Table 2.4: Optimal Cooperative Value Function W_{it} per country i for each period of time t in billions of 1990 USA dollars.

t	USA	Japan	EU	China	FSU	ROW	Total
1	10.36	3.04	11.47	0.79	0.82	40.10	66.60
2	23.70	9.99	26.21	2.11	2.56	80.57	145.17
3	46.27	22.43	56.22	4.11	5.10	114.70	248.85
4	74.97	38.79	93.26	7.22	8.48	155.73	378.46
5	110.47	56.52	136.58	11.06	12.79	215.05	542.49
6	150.04	77.48	185.37	15.53	17.40	281.58	727.43
7	193.44	101.33	241.07	20.62	22.53	392.50	971.51
8	242.66	125.29	302.21	26.31	28.05	433.63	1158.17
9	289.74	151.39	358.96	32.28	33.85	575.99	1442.23
10	337.33	175.75	419.29	38.32	39.61	709.60	1719.92
11	385.02	199.60	476.43	44.81	44.97	785.26	1936.12
12	432.90	222.13	530.72	50.77	50.39	892.96	2179.90
13	474.95	242.58	580.07	57.79	55.08	1015.22	2425.71
14	512.75	260.73	623.31	63.75	59.21	1126.00	2645.77
15	545.18	277.42	658.67	69.50	62.97	1208.34	2822.10
16	573.90	289.57	691.63	74.55	65.99	1275.03	2970.68
17	601.61	299.14	718.56	78.83	69.49	1386.34	3153.99
18	623.18	307.19	738.31	83.78	70.96	1379.83	3203.28
19	636.70	311.09	748.38	87.70	72.69	1507.28	3363.86
20	646.94	316.81	757.79	90.56	74.61	1414.53	3301.26
21	656.94	315.80	761.81	93.82	74.64	1462.77	3365.80
22	660.04	313.58	759.36	96.34	75.48	1496.71	3401.53
23	654.25	310.59	751.37	98.16	73.76	1518.74	3406.88
24	645.97	305.77	739.95	98.82	72.47	1537.81	3400.82
25	633.97	297.62	721.16	99.57	71.05	1557.01	3380.40
26	623.53	288.23	700.42	101.49	69.40	1540.37	3323.45
27	605.08	277.14	677.44	99.43	67.15	1515.21	3241.47
28	581.22	263.68	638.92	96.74	64.91	1480.54	3126.03
29	560.10	250.80	618.51	95.80	61.93	1421.17	3008.33
30	540.72	241.44	596.75	95.13	59.79	1356.44	2890.31
31	514.51	227.03	564.01	93.01	56.41	1352.03	2807.03
32	482.63	217.06	528.24	88.25	53.18	1322.74	2692.13
33	462.26	201.73	492.23	85.54	51.60	1216.81	2510.19
34	439.45	188.74	473.92	83.49	48.22	1152.18	2386.02
35	404.44	178.00	440.40	77.55	46.01	1092.96	2239.38
36	389.31	171.13	415.47	76.43	42.14	1010.63	2105.14
37	363.08	171.14	392.95	75.77	41.33	880.54	1924.83
38	356.62	162.54	386.85	74.48	40.17	831.40	1852.09
39	329.66	150.44	356.56	68.72	37.58	965.29	1908.27
40	315.54	147.68	345.07	69.52	37.49	804.88	1720.20
Total	17131.63	8168.56	19716.07	2628.62	1942.43	40506.65	90093.99

Table 2.5: Optimal non cooperative emissions e_{it}^N for each country at each period of time t in billion tons of carbon equivalent.

t	USA	Japan	EU	China	FSU	ROW	Total
0	1.370	0.292	0.872	0.805	1.066	3.430	7.835
1	1.693	0.329	1.081	0.946	1.193	4.811	10.055
2	1.736	0.309	1.060	0.993	1.241	5.214	10.554
3	1.807	0.349	1.124	1.020	1.270	5.379	10.952
4	1.864	0.370	1.156	1.095	1.300	5.484	11.272
5	1.933	0.361	1.175	1.165	1.353	5.681	11.670
6	1.987	0.357	1.184	1.227	1.375	5.846	11.977
7	2.050	0.385	1.234	1.300	1.411	6.444	12.827
8	2.134	0.371	1.289	1.369	1.440	6.129	12.734
9	2.193	0.405	1.303	1.444	1.494	6.999	13.839
10	2.265	0.416	1.364	1.515	1.549	7.610	14.722
11	2.341	0.426	1.403	1.604	1.580	7.759	15.115
12	2.438	0.436	1.445	1.662	1.632	8.174	15.790
13	2.516	0.448	1.488	1.800	1.668	8.724	16.647
14	2.594	0.461	1.528	1.895	1.700	9.227	17.407
15	2.661	0.481	1.552	1.996	1.742	9.600	18.035
16	2.733	0.481	1.595	2.083	1.766	9.927	18.588
17	2.849	0.491	1.653	2.172	1.872	10.590	19.628
18	2.939	0.502	1.690	2.307	1.877	10.608	19.926
19	3.029	0.513	1.728	2.441	1.953	11.500	21.167
20	3.098	0.539	1.764	2.529	2.035	11.078	21.046
21	3.204	0.534	1.807	2.669	2.057	11.616	21.890
22	3.301	0.541	1.851	2.814	2.148	12.139	22.796
23	3.369	0.561	1.892	2.959	2.143	12.652	23.578
24	3.450	0.583	1.941	3.089	2.183	13.221	24.469
25	3.537	0.595	1.980	3.257	2.244	13.865	25.480
26	3.660	0.611	2.029	3.495	2.315	14.348	26.460
27	3.745	0.621	2.078	3.610	2.368	14.822	27.247
28	3.799	0.614	2.047	3.710	2.427	15.264	27.862
29	3.877	0.610	2.119	3.895	2.451	15.555	28.508
30	3.975	0.644	2.186	4.107	2.526	15.818	29.257
31	4.039	0.637	2.205	4.288	2.546	16.614	30.332
32	4.060	0.673	2.206	4.369	2.574	17.245	31.128
33	4.158	0.646	2.181	4.533	2.689	17.220	31.430
34	4.235	0.631	2.276	4.742	2.695	17.521	32.102
35	4.171	0.632	2.245	4.712	2.759	17.829	32.350
36	4.294	0.687	2.261	4.977	2.672	17.792	32.685
37	4.200	0.758	2.223	5.199	2.786	16.265	31.434
38	4.187	0.688	2.210	5.185	2.711	13.391	28.374
39	4.094	0.620	2.107	4.880	2.686	19.557	33.945
40	3.890	0.663	2.042	5.231	2.856	16.000	30.683
Total	124.125	20.996	69.720	114.302	81.301	465.535	875.981

Table 2.6: Optimal Non Cooperative Value Function N_{it} for each country i for each period of time t in billions of 1990 USA dollars.

t	USA	Japan	EU	China	FSU	ROW	Total
1	5.06	5.84	6.77	0.47	0.60	7.93	26.68
2	19.42	11.18	26.00	1.89	2.37	31.62	92.50
3	43.10	25.93	56.80	4.29	5.23	71.06	206.42
4	75.34	45.25	98.79	7.66	9.14	125.68	361.89
5	115.92	65.95	151.23	12.26	14.11	196.02	555.51
6	164.68	91.64	213.45	17.58	20.00	281.70	789.07
7	219.67	122.34	283.64	24.08	26.76	381.76	1058.27
8	282.16	155.79	361.84	32.23	34.08	496.27	1362.38
9	349.47	191.70	446.51	40.71	42.21	621.86	1692.47
10	423.47	231.09	537.75	51.09	51.24	762.28	2056.94
11	501.90	271.61	633.71	60.63	60.39	914.52	2442.79
12	582.70	314.34	732.07	72.19	70.56	1078.52	2850.40
13	670.69	358.67	837.52	85.19	80.66	1253.13	3285.89
14	758.07	403.69	941.22	100.20	91.88	1432.80	3727.87
15	846.31	449.14	1048.52	114.32	102.46	1623.33	4184.10
16	939.33	496.08	1154.97	128.34	114.58	1822.98	4656.30
17	1033.52	541.15	1264.91	146.11	125.49	2025.46	5136.66
18	1127.86	586.75	1368.31	164.33	136.87	2237.54	5621.69
19	1222.51	632.46	1478.10	181.05	148.04	2451.76	6113.93
20	1308.22	677.36	1576.07	199.06	157.74	2667.92	6586.39
21	1401.98	718.26	1675.83	219.71	169.97	2883.85	7069.61
22	1494.40	764.77	1781.17	239.49	181.54	3114.14	7575.53
23	1581.51	803.59	1873.40	259.72	191.21	3332.86	8042.31
24	1672.01	841.28	1968.21	279.32	203.50	3564.32	8528.67
25	1757.22	881.46	2058.03	298.55	213.23	3786.39	8994.90
26	1839.20	917.55	2139.63	318.38	224.03	4011.68	9450.49
27	1907.95	951.24	2214.52	342.94	230.62	4230.27	9877.57
28	1990.59	979.13	2287.86	362.21	241.65	4449.01	10310.47
29	2055.34	1013.47	2364.14	385.94	250.03	4664.06	10733.01
30	2126.25	1037.72	2422.63	405.75	256.61	4869.33	11118.32
31	2195.01	1066.50	2484.03	429.62	266.21	5080.22	11521.61
32	2251.36	1085.61	2538.38	449.49	272.59	5282.73	11880.19
33	2308.94	1107.41	2590.56	472.21	282.14	5481.46	12242.74
34	2367.81	1132.14	2642.89	496.39	285.54	5680.21	12604.99
35	2416.26	1152.26	2679.95	518.21	292.52	5881.18	12940.40
36	2461.81	1161.56	2720.95	538.01	299.27	6065.00	13246.63
37	2514.89	1182.82	2752.02	560.59	303.99	6248.38	13562.70
38	2550.68	1189.11	2788.44	581.49	309.67	6431.11	13850.52
39	2585.63	1200.26	2808.86	607.00	314.70	6597.56	14114.04
40	2633.17	1211.49	2839.46	627.55	320.02	6785.05	14416.75
Total	52801.57	26075.77	60849.32	9836.43	6403.64	118923.05	274889.81

Table 2.7: Optimal stocks of pollutant cooperative s_t^W and non cooperative s_t^N and their differences for each period of time t in billion tons of carbon equivalent.

t	s_t^W	s_t^N	Difference
0	590.0000	590.0000	0.0000
1	544.5887	546.4247	1.83592
2	503.2917	507.3095	4.01784
3	465.7098	472.1477	6.43787
4	431.6674	440.5019	8.83448
5	400.5691	412.1310	11.56194
6	372.1760	386.6461	14.47010
7	346.5303	364.3270	17.79666
8	323.1425	343.9441	20.80159
9	301.9700	326.5193	24.54933
10	282.6366	311.5612	28.92458
11	265.1027	298.3557	33.25308
12	249.1796	287.0259	37.84635
13	234.5776	277.5828	43.00516
14	221.3860	269.7580	48.37202
15	209.4455	263.2723	53.82673
16	198.5148	257.9288	59.41401
17	188.5776	254.1118	65.53429
18	179.5473	250.9391	71.39189
19	171.2710	249.2960	78.02505
20	163.9237	247.6811	83.75741
21	157.1870	247.0574	89.87035
22	150.8389	247.3965	96.55765
23	145.2260	248.4863	103.26028
24	140.0085	250.3680	110.35947
25	135.2929	253.0903	117.79739
26	131.0180	256.5451	125.52711
27	127.2437	260.4724	133.22862
28	123.7602	264.6583	140.89812
29	120.5500	269.1092	148.55920
30	117.7626	273.9050	156.14239
31	115.1078	279.3397	164.23183
32	112.8148	285.0763	172.26153
33	110.6808	290.5929	179.91211
34	108.6996	296.2809	187.58129
35	106.9226	301.6991	194.77650
36	105.3961	306.9598	201.56373
37	103.9672	310.4921	206.52493
38	102.7126	310.6428	207.93016
39	101.6355	316.3512	214.71562
40	100.5178	318.2784	217.76052

2.6 Chapter summary and Extensions

In this chapter, we have developed a useful stochastic formulation which extends the stock pollutant control model developed by Germain et al. (2003). Our model lets to include through the random disturbance term, random elements not considered in the deterministic model. Moreover, our proposal lets to evaluate the magnitude of this effects by estimating, for example, the variance of the additive noise process. In principle we have assume independence for this process but we can also extend our work by considering some time series structures for the noise process.

Additionally, our example shows that the stochastic formulation produce consistent results in comparison to the deterministic model of reference, but simultaneously provides more flexibility than the former one. Note that the example proposed to illustrate our formulation is very close to the CLIMNEG model, which has been in fact analyzed from the deterministic point of view. So, in somehow we also extend this model to a stochastic setting. On the other hand, we want to remark that our real data based example is strongly driven by the original values taken at 1990 according to the ONU (1997).

Summarizing our results, for each country $i \in J$ and each period of time $t \in \mathcal{T}$ we obtain the following stock pollution, emissions and values functions for each model, $\{s_t^W\}$, $\{e_{it}^W\}$, $\{W_i(t, s_{t-1}^W)\}$ corresponding to the Pareto equilibrium of the Cooperative Model $(P1)$, and $\{s_t^N\}$, $\{e_{it}^N\}$, $\{N_i(t, s_{t-1}^N)\}$ corresponding to the Nash equilibrium of the Non-Cooperative Model $(P2)$. One might think an extension of our stochastic model by considering monetary transfers to induce cooperation, having in mind the significative differences between the optimal cooperative and non cooperative stock pollutant pointed out in our example, see for instance Figure 2.5. This is in fact research presented in 3.

Stochastic performance criteria based on bounds of probability that a particular country does not exceed a predefined total discounted cost, could be also considered, as an extension of this work and it is part of the results presented in 4.

From the infinite horizon optimization point of view, the present work is also related with developments in Hernández-Lerma and Romera (2001). Finally, further research could be done if we consider uncertainty about the random perturbation, say the variance of the i.i.d. sequence. We propose to estimate the parameter recursively and to include the estimation in the stochastic control problem.

Capítulo 3

The International Stock Pollutant Control: A Stochastic Formulation with Transfers

"Se mide la inteligencia de un individuo
por la cantidad de incertidumbre
que es capaz de soportar"

Immanuel Kant

Abstract

This chapter provides a stochastic dynamic game formulation of the economics of international environmental agreements on the transnational pollution control, when the environmental damage arises from stock pollutant that accumulates, for accumulating pollutants such as CO_2 in the atmosphere. To improve the non-cooperative equilibrium among countries, we propose a criteria to the minimization of the expected discounted total cost with monetary transfers between the countries involved as incentive to cooperation. Moreover, it considers the Stochastic Dynamic Games formulated as Markov Decision Processes, using tools of Stochastic Optimal Control and Stochastic Dynamic Programming, for solve Cooperative versus Non-cooperative stochastic dynamic Games. The performance of the proposed schemes is illustrated by a real data based example.

3.1 Introduction

In the framework of a deterministic cooperative game with a dynamic, multi-regional integrated assessment model, Eyckmans and Tulkens (2003) calculated the optimal path of abatement and aggregated discounted welfare for each region. They apply the transfer scheme advocated by Chander and Tulkens (1997) for the Climate Negotiation (CLIMNEG) World Simulation Model (abbreviated as CWSM) with six regions, the idea of surplus sharing is used for determining the transfer scheme, and they compute all possible partial agreement Nash equilibria. They found that allocation in the full cooperation lies in the core of the emission abatement game under this specific transfer scheme. Their CWSM derived from the seminal multi-region economy-climate Regional Integrated model of Climate and the Economy (RICE) of Nordhaus and Yang (1996).

The transfer schemes are based on a single year for assigning the permits or shares in the surplus. Such static schemes are also often observed in reality, e.g. the reduction targets in the ONU (1997) are designed as reduction compared to 1990 levels. These static schemes, however, do not take into account that the future growth paths of emissions are expected to diverge substantially between regions. This leads to assignments where historically large emitters obtain relatively large shares of the permits/surplus, while fast-growing developing countries, as China or India, obtain relatively small shares. This leads to increasing burdens on these developing countries to reduce theirs emissions; a notion brought forward by many developing countries in their argumentation on why they do not agree on any reduction targets in the ONU (1997).

The role of transfers in the analysis of self-enforcing International Environmental Agreements (IEA) was developed in Carraro et al. (2006). They propose transfers using internal and external financial resources for making welfare optimal agreements. To illustrate the relevance of their transfer scheme, they use a stylized integrated assessment simulation model of climate change to show how appropriate transfers may induce almost all countries into signing a self-enforcing climate treaty.

The studies by Germain et al. (2003) have addressed the issue of how many countries will be interested in signing an IEA with stock pollutant, adopting a cooperative game-theory approach. They extend the result established by Chander and Tulkens (1995) and (1997) for flow pollutants to the larger context of closed-loop (feedback) dynamic games with a stock pollutant. In this context, cooperation is negotiated at each period but financial transfers provide incentives to the countries that ensures the implementation of the grand coalition at each period. Their model thus yields a sequence of full cooperative international agreements so that full cooperation is also achieved in a dynamic setting with a stock pollutant.

Another paper related with this issue using a cooperative game-theory approach is Petrosjan and Zaccour (2003). However, in this paper the authors assume that all the countries decide to cooperate at the initial time-consistent decomposition of each player's total cost, as given by Shapley value, so that the countries stick at each moment to the full cooperative solution agreed at initial time, supposing that the global allocation problem has been solved.

The main purpose of this chapter is to suggest a stochastic dynamic game formulation for the Stock Pollutant Control model closer to reality, because we consider random factors and not only economic factors. This model proposed is directly linked with the Kyoto or post-Kyoto agreement mechanisms. We use financial transfers as additional elements in the game for the design of international agreements that achieve global optimality in stock pollutant control problems.

The stochastic formulation for this Stock Pollutant Control Model involve the use of Stochastic Dynamic Programming with discrete and finite planing horizon, for searching both cooperatives and non cooperatives equilibria with transfers. Stochastic optimization problems should be solved by Stochastic Dynamic Programming Techniques, i.e. see Bertsekas (2000) and Birge and Louveaux (1997).

Stochastic Programming is considered by Dechert and O'Donnell (2006) in a particular application that explore some fundamental issues of the optimal level of pollution in a lake with competing uses, they show how the model can be interpreted as an open loop dynamic game, where the control variables are the levels of phosphorus discharged into the watershed of the lake, the state of the system is the accumulated level of phosphorus in the lake and the random shock (a multiplicative noise factor on the control variables of the players) is the rainfall that washes the phosphorus in the lake.

This Chapter is organized as follows: In Section 3.2, we define the monetary transfers to ensure that each country is not worse off when it participates and we report of non cooperative model with monetary transfer in each period of time for each country, with necessary definitions and results. In Section 3.3, we present an algorithm which solves the problem with transfers of minimize the expected discounted total cost, the optimal action sets, and optimal policies for a finite horizon model for each period of time and for all the countries. In Section 3.4, we present a numerical example based on real scenarios borrowed from the work by Eyckmans and Tulkens (2003) and Casas and Romera (2009a), as in Chapter 2. In Section 3.5, we present some conclusions and extensions of this chapter.

3.2 Transfers Definition

At the international optimum ($P1$), and contrary to what happens at the Nash equilibrium ($P2$), each country takes account of the impact of its pollution on the environment of all other countries. Therefore, from a collective point of view, the optimum is better than the Nash equilibrium. Nothing ensures that this is also true at the individual level. Indeed, countries being different, it is possible that some country i at some period of time t is better off at the non cooperative equilibrium than the optimum, so that cooperation is not profitable for this country, at least at time t. The same can occur for subsets of countries (i.e. coalitions) in the sense that, by limiting cooperation to such coalitions, the members of the latter could be better off than at the international optimum.

In a deterministic dynamic programming framework, Germain et al. (2003) propose a mechanism of financial transfers between countries that can make of them interested in achieving the international optimum at all periods t (*individual rationality*). This mechanism has the additional property that no subgroup of countries has ever an incentive to form a coalition and enact an optimum for itself only (*coalition rationality*). In the present Chapter 3, our aim is to apply the mechanism proposed by Germain et al. (2003) to the climate change with a stochastic model.

At each time period t one could take the non cooperative Nash equilibrium from t onwards as such point of reference, and determine the transfers accordingly. However, one should not neglect the fact that countries know that later on, thanks to the cooperative transfers to which they will have access, they will be better off than at the non cooperative Nash equilibrium. Hence, a better point of reference at period of time t is non cooperation at time t, followed by cooperation afterwards. We are thus aiming at a cooperative international optimum.

3.2.1 Transfers at final period T

We start by determining which transfers yield for all countries when they cooperate in the last period $t = T$, of the finite horizon $\mathcal{T}$, for any level of the stock of pollution s_{T-1} inherited from the past.

In the non cooperative equilibrium the countries are supposed to solve problem $(P2)$, the country i's expected total cost at period final T is then

$$N_i(T, s_{T-1}^N) = c_i(e_{iT}^N) + d_i(s_T^N),$$

where $e_{iT}^N = (e_{1T}^N, e_{2T}^N, \ldots, e_{nT}^N)$ denotes the vector of emissions equilibrium level of each countries and s_T^N denotes the resulting stock of pollutant given by

$$s_T^N = [1-\delta]s_{T-1}^N + \sum_{i=1}^n e_{iT}^N,$$

where s_{T-1}^N is the inherited stock of pollutant at the begin of period T.

If countries cooperate, they jointly solve $(P1)$. The country i's expected total cost at final period T is

$$W_i(T, s_T^W) = c_i(e_{iT}^W) + d_i(s_T^W),$$

where $e_{iT}^W = (e_{1T}^W, e_{2T}^W, \ldots, e_{nT}^W)$ is the vector of optimal emission level (policy) and s_T^W is the optimal stock of pollutant at final period T, given by

$$s_T^W = [1-\delta]s_{T-1} + \sum_{i=1}^n e_{iT}^W,$$

where s is the inherited stock of pollutant at the begin of period T.

Let defines $W(T, s_T^W)$ and $N(T, s_T^N)$ as the expected total cost cooperative and non cooperative, respectively, at final period of time T. By definition of the optimum, one verifies that

$$W(T, s) \equiv \sum_{i=1}^n W_i(T, s) \quad \leq \quad \sum_{i=1}^n N_i(T, s) \equiv N(T, s). \tag{3.1}$$

The difference $W(T, s) - N(T, s)$, between the two sides of this inequality (3.1) measures the *ecological surplus* resulting from international cooperation.

However, the inequality (3.1) is not sufficient to ensure cooperation. Indeed, if $\exists i \in J$ such that $W_i(T,s) > N_i(T,s)$, then country i will not cooperate without financial compensation for higher cost it incurs. Since stochastic dynamic programming reduces the choice of emissions to one period at the time, one can use the transfers formula in a static framework.

Following the transfers formula proposed by Chander and Tulkens (1997) in a static framework, we can use this transfers formula in our stochastic dynamic framework at final period T

$$\theta_i(T,s) \quad = \quad -[W_i(T,s) - N_i(T,s)] + \mu_{iT}[W(T,s) - N(T,s)], \tag{3.2}$$

with

$$\sum_{i=1}^{n} \theta_i(T,s) = 0.$$

The transfer (3.2) is < 0 if received and > 0 if paid to country i at period T, and it satisfies that $\mu_{iT} \in]0;1[$, $\forall i \in J$, and

$$\sum_{i=1}^{n} \mu_{iT} = 1.$$

The fact that μ_{iT} cannot be equal to 0 ensure that country i will benefit from cooperation if $W_i(T,s) < N_i(T,s)$. The fact that μ_{iT} cannot be equal to 1 exclude that country i monopolizes all the gains of cooperation.

Then country i's total cost including transfers at final period T becomes

$$\tilde{W}_i(T,s) = W_i(T,s) + \theta_i(T,s). \tag{3.3}$$

The cooperation with transfers is *individually rational* at final period of time T, in the sense that each country have interest to participate whatever the inherited stock of pollutant s, since by construction

$$\tilde{W}_i(T,s) - N_i(T,s) = \mu_{i,T}\left[W(T,s) - N(T,s)\right] \leq 0, \quad \forall i \in J.$$

3.2.2 Transfers at period $T-1$

The countries know that, whatever they do al period $T-1$, financial transfers exist defined by (3.2), that make the international optimum (cooperative) at period T preferable for each of them with respect to the non cooperative equilibrium. Let us assume that these transfers induce cooperation, following Chander and Tulkens (1997), one could indeed obtain the cooperative optimum with transfers as an equilibrium, called *ratio equilibrium*, and that countries therefore expect, at period of time $T-1$, that they will cooperate in period T. This is the rational expectations assumption. The problem we wish to consider now is wether under assumption transfers can be designed that make the countries interested to cooperate at period $T-1$ as well.

In absence of cooperation at $T-1$, each country $i \in J$ minimizes its own expected discounted total cost over two periods $T-1$ and T, expecting cooperation and transfers at period T. Thus, given the emissions of the other countries, the country i solves problem $(P2)$ for $t = T-1$ with expected transfers at period T.

There are n problems to solve at period $T-1$, each country $i \in J$ solves the following problem

$$\min_{e_{i,T-1}} \quad \mathbb{E}\left[c_i(e_{i,T-1}) + d_i(s_{T-1}) + \beta \tilde{W}_i(T, s_{T-1})\right] \tag{3.4}$$

$$\text{s.t.} \quad s_{T-1} = (1-\delta)s_{T-2} + \sum_{i=1}^{n} e_{i,T-1} + \xi_{T-1}$$
$$e_{i,T-1} \geq 0, \quad \forall i \in J.$$

Since the expected value functions $\tilde{W}_i(T, s_{T-1})$ contain transfers that sum up to zero, convexity of the cost functions c_i and damages function d_i ensure that the objectives in (3.4) are convex, unlike what happens for the period T.

This yields an equilibrium characterized at period $T-1$ by emissions level $e^V_{i,T-1}$ as functions of initial stock s at period $T-1$. The expected value functions $V_i(T-1, s)$ denotes country i's expected discounted non cooperative equilibrium costs including future transfers at final period T.

The *expected value functions* V_i, according to Bellman's principle of optimality

$$V_i(T-1, s) = c_i(e^V_{i,T-1}) + d_i(s^V_{T-1}) + \beta \tilde{W}_i(T, s^V_{T-1}), \quad \forall i \in J, \tag{3.5}$$

where

$$s^V_{T-1} = (1-\delta)s_{T-2} + \sum_{i=1}^{n} e^V_{i,T-1},$$

denotes country i's expected discounted equilibrium costs. We will call this equilibrium the non cooperative equilibrium with transfers at period $T-1$.

In the case where all countries cooperate, they solve problem $(P1)$ for $t = T-1$. Optimal levels of emissions and of the resulting stock of pollutant are denoted by $e^W_{i,T-1}$ and s^W_{T-1}, respectively, both are function of the initial stock s at period $T-1$. This yields

$$W_i(T-1, s) = c_i(e^W_{i,T-1}) + d_i(s^W_{T-1}) + \beta \tilde{W}_i(T, s^W_{T-1}), \quad \forall i \in J, \tag{3.6}$$

which is country i's part in the optimal expected total discounted costs, taking into account the transfers and the resulting cooperation expected at final period T.

As in period T, see (3.1), one verifies that

$$W(T-1, s) \equiv \sum_{i=1}^{n} W_i(T-1, s) \leq \sum_{i=1}^{n} V_i(T-1, s) \equiv V(T-1, s). \tag{3.7}$$

The difference $W(T-1, S) - V(T-1, s)$, between the two sides of this inequality (3.7) measures the *ecological surplus* induced by extending from international cooperation to period $T-1$, with respect to alternative scenario where cooperation is limited to period T.

However, (3.7) is again not sufficient to induce cooperation at time $T-1$, if exist $i \in J$ such that $W_i(T-1, S) > V_i(T-1, s)$, then country i will not want to extend cooperation to period $T-1$ without financial compensation.

To induce country i to participate at period $T-1$, let

$$\theta_i(T-1, s) = -[W_i(T-1, s) - V_i(T-1, s)] + \mu_{i,T-1}[W(T-1, s) - V(T-1, s)], \tag{3.8}$$

be the transfer paid or received by country i at period $T-1$ where $\mu_{i,T-1} \in]0, 1[$, for all $i \in J$ and

$$\sum_{i=1}^{n} \mu_{i,T-1} = 1 \quad \text{and} \quad \sum_{i=1}^{n} \theta_i(T-1, s) = 0.$$

Then country i's expected total cost including transfers becomes

$$\tilde{W}_i(T-1, s) = W_i(T-1, s) + \theta_i(T-1, s).$$

It is clear that this transfers defined by (3.8) make cooperation individually rational at period $T-1$, whatever the inherited stock of pollutant s.

3.2.3 Transfers at period t

We can be repeated the preceding analysis for all earlier periods. The final result will be that the countries cooperate in each period t. This determines the emissions levels in each period for each countries, and also the trajectory of the stock of pollutant, given its initial value s_0. In turn this trajectory determines the expected value functions V_i, W_i and $\tilde{W}_i$, and therefore also the value of the transfers θ_i.

There are n problems to solve, one for each country i, named $(P3)$

$$\begin{aligned}(P3) \qquad & \min_{e_{it}} \qquad \mathbb{E}\left[(c_i(e_{it}) + d_i(s_t)) + \beta\tilde{W}_i(t+1, s_t)\right] \\ & \text{s.t.} \qquad s_t = (1-\delta)s_{t-1} + \sum_{i=1}^{n} e_{it} + \xi_t \\ & \qquad\qquad e_{it} \geq 0 \quad \forall t \in \mathcal{T}, \quad \forall i \in J \\ & \qquad\qquad s_0 > 0\end{aligned}$$

This yields an equilibrium characterized at period t by emissions level e_{it}^V as functions of initial stock s_0 at each period t.

The value functions $V_i(t,s)$ denotes the country i's expected discounted non cooperative equilibrium costs including transfers at $t+1$ period

$$V_i(t, s_t) = c_i(e_{it}^V) + d_i(s_t) + \beta\tilde{W}_i(t+1, s_t), \qquad \forall i \in J,$$

where the transfer paid or received by country i at period t

$$\theta_i(t,s) = -[W_i(t,s) - V_i(t,s)] + \mu_{i,t}\left[\sum_{i=1}^{n} W_i(t,s) - \sum_{i=1}^{n} V_i(t,s)\right]. \tag{3.9}$$

be the transfer paid or received by country i at period t where $\mu_{i,t} \in]0,1[$, for all $i \in J$,

$$\sum_{i=1}^{n} \mu_{i,t} = 1 \quad \text{and} \quad \sum_{i=1}^{n} \theta_i(t,s) = 0.$$

Then country i's expected total cost including transfers becomes

$$\tilde{W}_i(t,s) = W_i(t,s) + \theta_i(t,s).$$

It is clear that this transfers (3.9) make cooperation individually rational at period t, whatever the inherited stock of pollutant s.

3.2.4 Alternative problem with transfers

Following the same technique as in Casas and Romera (2009a) and in 2, subsection of Non-Cooperative Alternative Problem, and bearing in mind the explicit recursive expression (2.2) obtained for the stock pollutant s_t, we have to solve for each country $i \in J$ the alternative problem with future transfers, as following

$$\min_{\{e_{it}\}} \quad \mathbb{E}\left[\left[c_i(e_{it}) + \tilde{d}_i(s_0, e_{it}, \xi_t)\right] + \beta \tilde{W}_i(t+1, s_t)\right]$$

$$\begin{array}{lrcl} \text{s.t.} & M_i e_{it} & = & b_i + \xi \\ & e_{it} & \geq & 0 \quad \forall t \in \mathcal{T}, \quad \forall i \in J \\ & s_0 & > & 0 \end{array}$$

where

$$\begin{array}{rcl} e_i' & = & (e_{i1}; e_{i2}; \cdots ; e_{i,T-1}) \\ b_i' & = & (b_{i1}; b_{i2}; \cdots ; b_{i,T-1}) \\ b_{it} & = & -(1-\delta)^t s_0 - \displaystyle\sum_{\tau=1}^{t} \sum_{j \neq i}^{n} (1-\delta)^{t-\tau} e_{j\tau} \\ \xi' & = & \big(\xi_1; (1-\delta)\xi_1 + \xi_2; \ldots ; \ldots ; (1-\delta)^{T-1}\xi_1 + (1-\delta)^{T-2}\xi_2 + \cdots + \xi_T\big) \end{array}$$

The matrix M_i is a square matrix, lower triangular, of order T, with ones in the principal diagonal. The vector b and the random disturbance ξ have order T. The structure of the matrix M_i is as follows

$$M_i = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \cdots & 0 \\ (1-\delta) & 1 & 0 & 0 & 0 & \cdots & 0 \\ (1-\delta)^2 & (1-\delta) & 1 & 0 & 0 & \cdots & 0 \\ (1-\delta)^3 & (1-\delta)^2 & (1-\delta) & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & 0 \\ (1-\delta)^{t-1} & (1-\delta)^{t-2} & (1-\delta)^{t-3} & \cdots & \cdots & (1-\delta) & 1 \end{pmatrix}$$

then we can may obtain the inverse matrix of the matrix B_i, which is quasi diagonal

$$M_i^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \cdots & 0 \\ -(1-\delta) & 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & -(1-\delta) & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & -(1-\delta) & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & \cdots & -(1-\delta) & 1 \end{pmatrix}$$

As in the cooperative model solution ($P1$), of Section 2.3, and in the non cooperative model solution ($P2$) of Section 2.4, by using the development presented in this section one can find the parameters e_{it}^V of optimal emissions for each country $i \in J$, and one obtains the stock levels of contamination s_t^V at each period of time $t \in \mathcal{T}$.

In the infinite horizon case ($T = \infty$), the backward reasoning considered above applies no more. However, we can consider the stationary solution by taking advantage of the fact that the cost functions c_i and damage functions d_i as well as the sharing parameters μ_i do not depend directly on time. The functional forms of the solutions thus only vary in time through the varying stock of pollutant s, see Appendix.

3.3 The Algorithm

The aim of this section is to describe a numerical algorithm that calculates the expected value functions and the transfers when the cost functions c_i and d_i are convex. It is written for finite horizon problems, which is appropriate for most practical applications. The emissions and stock trajectories associated with the international cooperative optimum ($P1$) can be calculated by non linear stochastic programming techniques. The transfers are more difficult to calculate, because they make use of values calculated at non cooperative equilibrium ($P2$), so that one must proceed by backward induction. The basic idea is to construct explicit approximation for the surfaces or value functions $\tilde{W}_i(t, s_{t-1})$, for all $i \in J$, as polynomial functions of s_{t-1}, by using classical regression.

The first step of the algorithm is to solve the cooperative problem ($P1$) associated with the international optimum. This is done by using non linear stochastic programming techniques and leads to the optimal trajectories of the abatement rates e_{it}^W for all $i \in J$ and for all $t \in \mathcal{T}$ and of the CO_2 stock s_t^W with $t \in \mathcal{T}$.

The second step consists of computing the financial transfers, and this supposes to solve the stochastic dynamic programming problems associated to the non cooperative equilibrium. As the algorithm proceeds backwards, this is first done the final time T using the system of first order conditions associated to non cooperative problem $(P2)$.

Hence one obtains the non cooperative abatement rates e_{jT}^N, the transfers $\theta_i(s,T)$ and the expected discounted total costs (transfers included) $\tilde{W}_i(T,s)$ for all $t \in \mathcal{T}$ as functions of the CO_2 stock s_T inherited at the beginning of time T.

The resolution for period T of the dynamic programming problems $(P2)$ is repeated for a set $\mathcal{S}_T$ of given values of the inherited CO_2 stock. This set is chosen to be representative of the interval of possible values of s_T. Once the values $\tilde{W}_i(T,s)$ for all $i \in J$ have been calculated on the set $\mathcal{S}_T$, the value functions $\tilde{W}_{iT}$ are written as polynomials of s_T and regressed on the set $\mathcal{S}_T$. So that the functions $\tilde{W}_{iT}$ are approximated whit explicit analytical functions of s_T.

The step 2 of the algorithm is repeated for period of time $T-1$, $T-2$, ... until period 1. At each period, one calculates the non cooperative equilibrium by solving the system of first order conditions associated to problems with monetary transfers $(P3)$. To do so at time t, the algorithm makes use of the polynomials regressed for period of time $t+1$.

Once the value functions $\tilde{W}_{it}$ are known as functions of s_t for all times of the planning period $t \in \mathcal{T}$, the algorithm performs its third step, i.e. the computation of the actual values of these value functions and of the transfers for all regions all along the optimal trajectory $e_1^W, e_2^W, ..., e_T^W$ calculated at first step.

3.3.1 Statement of the Algorithm

The algorithm consists of several steps

Step 1

The algorithm starts by solving the stochastic optimization cooperative problem $(P1)$

associated with the international optimum. The first order conditions are

$$c_i'(e_{it}^W) + \sum_{\tau=t}^{T} \beta^{\tau-t}[1-\delta]^{\tau-t} \sum_{j=1}^{n} d_j'(s_\tau^W) = 0 \qquad \forall i \in J; \forall t \in \mathcal{T} \tag{3.10}$$

$$s_t^W = [1-\delta]s + \sum_{i=1}^{n} e_{it}^W \qquad \forall t \in \mathcal{T} \tag{3.11}$$

where c_i' and d_j' are the derivatives of functions c_i and d_j respectively. Solving this $T(n+1)$ equations yields the optimal values of emissions e_{it}^W and the stock of pollutant s_t^W for all periods of time $t \in \mathcal{T}$.

Step 2a

To calculate the transfers, one must first solve the stochastic dynamic programming problems associated with the non cooperative equilibrium. the algorithm proceeds backwards, starting from the last period. At final time T, given the inherited stock of pollutant s, the non cooperative equilibrium, which coincides with the Nash equilibrium at last period, has to satisfy the conditions

$$c_i'(e_{iT}^N) + d_i'\left([1-\delta]s + \sum_{j=1}^{n} e_{jT}^N\right) = 0 \qquad \forall i \in J \tag{3.12}$$

Once this system of equations has been solved for the variables e_{jT}^N, and knowing by step 1 the optimal emissions at time T, is possible to calculate the transfers $\theta_i(s,T)$ using the transfers of the form

$$\tilde{\theta}_i(s) = -[c_i(e_i^W) - c_i(e_i^V)] + \tilde{\mu}_i(s^W) \sum_{j=1}^{n} [c_j(e_j^W) - c_j(e_j^V)] \tag{3.13}$$

with

$$\tilde{\mu}_i(s_T^W) = \frac{d_i'(s_T^W)}{\sum_{j=1}^{n} d_j'(s_T^W)} \tag{3.14}$$

as well as the value functions $\tilde{W}_i(T,S)$ defined by

$$\tilde{W}_i(s) = W_i(s) + \tilde{\theta}_i(s), \quad \forall i \in J. \tag{3.15}$$

The **Step 2a** is done for a set Ω_T of given values of the inherited stock of pollutant s. This set is chosen to be representative of the interval of possible values of s. This interval is bounded below by the value of s that would be obtained with zero emissions during periods $1, 2, \ldots, T-1$ given s_0, and bounded above by the value of s that would be obtained with maximum emissions during the same period of time.

Step 2b

We now assume that the total costs with transfers at time T have the following polynomial form

$$\tilde{W}_i(T, s) = k_{iT,m}s^m + k_{iT,m-1}s^{m-1} + ... + k_{iT,0} \qquad \forall i \in J$$

where m is the order of the polynomial chosen so that the fit is good enough. To identify the parameters $k_{iT,m}, k_{iT,m-1}, ..., k_{iT,0}$, we use an ordinary least square regression method implemented in the software Matlab, so that the functions $\tilde{W}_i(T, s)$ are now approximated with explicit analytical functions of s.

The Step 2a and 2b are then repeated for period $T-1$. Given (3.4) and the inherited stock of pollution s, the first order conditions associated to the non cooperative equilibrium $\forall i \in J$ are

$$c_i'(e_{it}^V) + d_i'\left([1-\delta]s + \sum_{j=1}^{n} e_{j,T-1}^V\right) + \beta\tilde{W}_i'\left(\left[[1-\delta]s + \sum_{j=1}^{n} e_{j,T-1}^V\right], T\right) = 0 \qquad (3.16)$$

Once this system of equations has been solved for the $e_{j,T-1}^V$, knowing the optimal emissions at time $T-1$ by Step 1, and the value functions $\tilde{W}_i(s, T)$ by applications of Steps 2a and 2b for the period of time T, it is possible to calculate the transfers $\theta_i(s, T-1)$ using (3.13)-(3.14) and the value functions $\tilde{W}_i(s, T-1)$ using (3.15). This is done for a set Ω_{T-1} of given values of the inherited stock of pollutant s, so that by regression of the calculated values of the value functions $\tilde{W}_i(s, T-1)$ on Ω_{T-1}, the value functions $\tilde{W}_i(s, T-1)$ can be approximated with explicit analytical functions of s.

The algorithm continues backwards by repeating the **Steps 2a** and **Steps 2b** until the first time period $t = 1$.

Step 3

Once the value functions $\tilde{W}_i$ are know as functions of S for all time of the planning period $\mathcal{T} = \{1, 2, ..., T\}$, the algorithm calculates the actual values of these value functions and of the transfers for all countries all along the optimal trajectory calculated at first Step.

3.4 A numerical Example

In this section, we show some numerical results obtained of solver the problem ($P3$), in a real scenario. The simulations are made for a time horizon of 100 years, we give the results only up to 2030, in order to avoid boundary problems. All computations were made by use of the software Matlab 7.3.0 (R2006b). Thus, we have developed specific code for our example.

The cost and damage functions used in this case are nonlinear and the arguments of these functions are selected according climate and economic principles. The temperature change equation is taken from the climate economy model RICE (Regional Integrated model of Climate and the Economy), as well as most of the parameter values and all basic data on GDP, population, capital stock, carbon emissions and concentration and global mean temperature. A complete overview of the equations and parameter values of the Climate Negotiation (CLIMNEG) World Simulation Model (abbreviated as CWSM) can be found in Eyckmans and Tulkens (2003) and Casas and Romera (2009a). The division of the world is the same as in the RICE model. There are 6 countries or regions: USA, Japan, European Union (EU), China, Former Soviet Union (FSU) and Rest of the World (ROW). The time is divided in years, the initial period (period $t = 0$) refers to year 1990.

Let consider Tables (2.4) and (2.6) of Chapter 2 which show optimal cooperative and non cooperative value functions respectively, for each country and for each period of time. After the first three rows we conclude that countries like USA or ROW have not incentive to cooperate, and the other hand Japan has from the beginning interest in cooperation. Thus, without any modification of the initial game the most rational behavior of countries will probably be the non cooperation one, and the Nash equilibrium will result in the optimal non-cooperative stock pollutant $\{s_t^N\}$. It means that the stock pollutant will be international optimum $\{s_t^W\}$.

Table (3.1) and Figure (3.4) show the the profile of financial transfers received and given θ_{it} between countries $i \in J$ for each period of time $t \in \mathcal{T}$. A negative transfer is a transfer received, and a positive transfer is a transfer given by the country i at period t. Note that

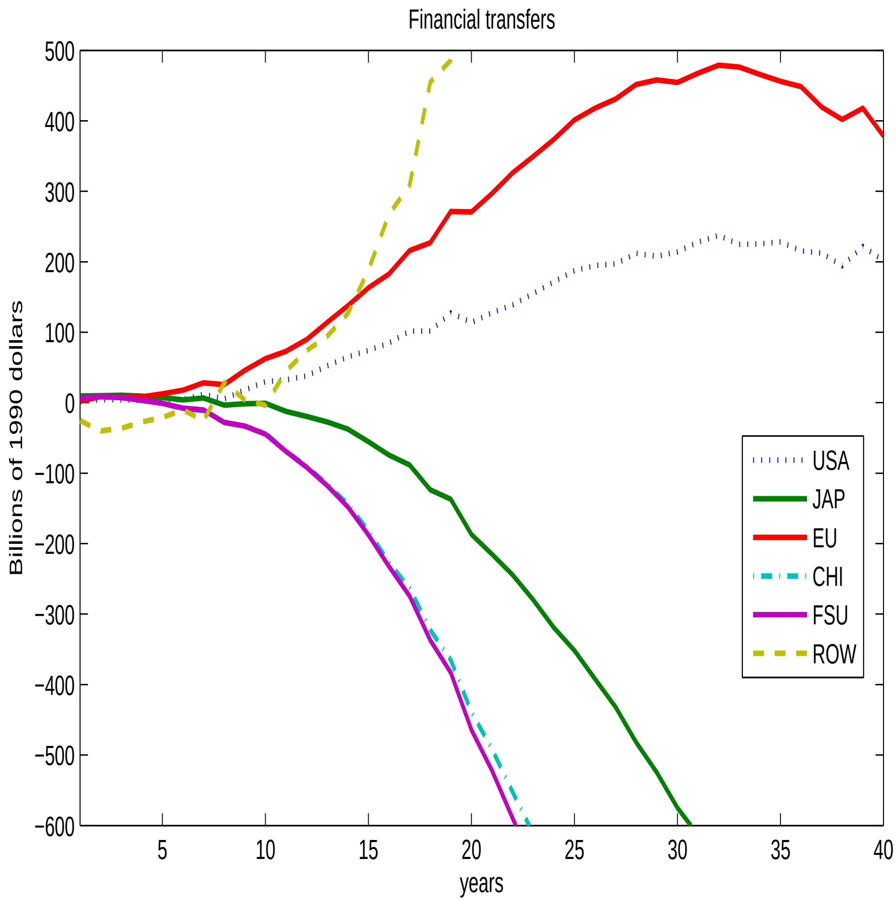

Figure 3.1: Financial transfers θ_{it} per country i for each period of time t in billions of 1990 USA dollars.

ROW, EU and USA pay at each period a transfer to China, Japan and FSU to induce these regions to cooperate.

Table (3.2) shows the optimal cooperative value function $\tilde{W}_{it}$ with transfers θ_{it} per country i for each period of time t. Let compare these results with the optimal cooperative value function W_{it} in Table (2.4) and N_{it} in Table (2.6) of Chapter 2. As it is expected the total value function is the same because the total sum od the transfers is equal to zero. Let compare the marginal total value function per country under the two cooperative models summarized in these Tables. We observe that USA, EU and ROW increase these values when transfers are considered in comparison to the basic cooperative game, and Japan, China and FSU are net receivers of transfers.

Table (3.3) resumes all the results related to the marginal optimal value function per country under the cooperative, non cooperative and cooperative with transfers games. Note that the distribution of the total value function in the model with transfers, the last column in Table (3.3), provides net given countries like USA, EU and ROW. This is necessary in order to achieve the international optimum $\{s_t^W\}$ which is the same as in the cooperative model ($P1$) of Chapter 2. This is fact a weakness of this proposal, the cooperative model with transfers ($P3$). A question arises, *Are there incentives enough to ensure that these countries will be able to accept that solution?*.

Table 3.1: Transfers received or given θ_{it} per country i for each period of time t in billions of 1990 USA dollars.

t	USA	Japan	EU	China	FSU	ROW	Total
1	1,344	9,448	1,950	6,337	6,438	-25,518	0
2	4,502	9,972	8,562	8,561	8,582	-40,179	0
3	3,897	10,566	7,657	7,252	7,199	-36,570	0
4	3,136	9,226	8,295	3,207	3,422	-27,286	0
5	3,280	7,255	12,482	-0,970	-0,847	-21,200	0
6	4,370	3,886	17,805	-8,220	-7,681	-10,160	0
7	11,772	6,548	28,113	-10,999	-10,233	-25,201	0
8	5,460	-3,537	25,595	-28,118	-28,005	28,605	0
9	18,018	-1,396	45,842	-33,272	-33,347	4,155	0
10	29,970	-0,824	62,288	-43,405	-44,534	-3,495	0
11	32,436	-12,435	72,837	-68,624	-69,029	44,814	0
12	38,045	-19,538	89,594	-90,332	-91,572	73,802	0
13	52,374	-27,270	114,094	-115,961	-117,777	94,539	0
14	64,968	-37,387	137,561	-143,903	-147,685	126,446	0
15	74,123	-55,283	162,851	-182,176	-187,505	187,990	0
16	84,494	-74,426	182,412	-227,146	-232,343	267,010	0
17	101,462	-88,426	215,900	-263,159	-274,450	308,674	0
18	101,608	-123,503	226,931	-322,520	-337,154	454,638	0
19	127,461	-136,978	271,382	-365,000	-382,995	486,131	0
20	113,754	-186,970	270,760	-439,026	-464,389	705,871	0
21	127,734	-214,839	296,716	-491,414	-521,973	803,776	0
22	138,693	-244,477	326,143	-552,522	-589,603	921,766	0
23	154,696	-279,573	349,452	-611,008	-655,119	1041,551	0
24	171,396	-319,129	373,623	-674,143	-723,614	1171,865	0
25	187,499	-351,904	401,119	-736,773	-793,565	1293,623	0
26	194,499	-391,849	418,039	-804,2897	-866,545	1450,145	0
27	196,856	-431,922	431,069	-862,506	-942,539	1609,042	0
28	211,961	-481,956	451,534	-931,934	-1020,664	1771,060	0
29	207,796	-524,782	458,192	-997,308	-1099,344	1955,446	0
30	214,195	-575,060	454,540	-1060,708	-1174,514	2141,547	0
31	228,066	-612,957	467,591	-1115,824	-1242,638	2275,761	0
32	237,386	-662,788	478,792	-1170,096	-1311,934	2428,640	0
33	224,590	-716,412	476,237	-1235,415	-1391,554	2642,554	0
34	225,200	-759,759	465,807	-1290,270	-1465,844	2824,866	0
35	228,321	-809,238	456,043	-1342,846	-1536,995	3004,715	0
36	215,579	-866,480	448,571	-1395,338	-1599,790	3197,458	0
37	212,161	-927,964	419,422	-1454,819	-1676,990	3428,191	0
38	194,319	-973,168	401,847	-1492,729	-1730,234	3599,965	0
39	221,677	-984,480	418,007	-1496,011	-1757,176	3597,984	0
40	201,541	-1052,285	378,299	-1558,071	-1833,559	3864,077	0
Total	4870,641	-12892,096	10333,951	-23591,495	-26338,098	47617,097	0

Table 3.2: Optimal Cooperative Value Function $\tilde{W}_{it}$ with transfers θ_{it} per country i for each period of time t in billions of 1990 USA dollars.

t	USA	Japan	EU	China	FSU	ROW	Total
1	11,713	12,493	13,424	7,128	7,260	14,587	66.607
2	28,207	19,968	34,779	10,677	11,149	40,399	145.178
3	50,170	33,001	63,877	11,363	12,306	78,134	248.851
4	78,112	48,018	101,561	10,428	11,903	128,446	378.469
5	113,754	63,781	149,067	10,096	11,943	193,851	542.492
6	154,415	81,375	203,179	7,314	9,726	271,426	727.435
7	205,217	107,883	269,183	9,623	12,303	367,301	971.510
8	248,125	121,759	327,812	-1,806	0,048	462,235	1158.172
9	307,766	149,995	404,808	-0,990	0,507	580,153	1442.238
10	367,301	174,926	481,586	-5,076	-4,924	706,112	1719.925
11	417,462	187,171	549,274	-23,807	-24,051	830,075	1936.124
12	470,951	202,598	620,323	-39,558	-41,182	966,771	2179.902
13	527,332	215,311	694,166	-58,167	-62,694	1109,769	2425.717
14	577,722	223,345	760,871	-80,148	-88,469	1252,451	2645.772
15	619,311	222,145	821,521	-112,676	-124,533	1396,334	2822.102
16	658,402	215,144	874,042	-152,595	-166,349	1542,044	2970.687
17	703,078	210,715	934,467	-184,326	-204,952	1695,016	3153.998
18	724,793	183,688	965,247	-238,731	-266,190	1834,475	3203.282
19	764,166	174,115	1019,765	-277,295	-310,298	1993,415	3363.869
20	760,703	129,847	1028,551	-348,459	-389,777	2120,401	3301.268
21	784,681	100,961	1058,533	-397,593	-447,326	2266,549	3365.805
22	798,736	69,107	1085,503	-456,175	-514,123	2418,482	3401.530
23	808,947	31,018	1100,829	-512,848	-581,356	2560,294	3406.884
24	817,373	-13,358	1113,577	-575,314	-651,136	2709,680	3400.821
25	821,474	-54,283	1122,285	-637,198	-722,511	2850,642	3380.409
26	818,029	-103,617	1118,464	-702,792	-797,140	2990,516	3323.459
27	801,938	-154,773	1108,514	-763,070	-875,387	3124,254	3241.477
28	793,182	-218,269	1090,459	-835,192	-955,750	3251,602	3126.033
29	767,898	-273,973	1076,702	-901,501	-1037,413	3376,618	3008.332
30	754,921	-333,611	1051,298	-965,577	-1114,717	3497,996	2890.311
31	742,584	-385,926	1031,607	-1022,805	-1186,220	3627,795	2807.036
32	720,020	-445,725	1007,040	-1081,844	-1258,747	3751,389	2692.133
33	686,851	-514,676	968,473	-1149,871	-1339,950	3859,373	2510.199
34	664,655	-571,019	939,731	-1206,772	-1417,620	3977,051	2386.026
35	632,763	-631,237	896,446	-1265,289	-1490,979	4097,682	2239.386
36	604,897	-695,347	864,043	-1318,900	-1557,640	4208,091	2105.144
37	575,247	-756,824	812,375	-1379,047	-1635,655	4308,736	1924.833
38	550,947	-810,619	788,701	-1418,248	-1690,059	4431,372	1852.094
39	551,345	-834,034	774,573	-1427,287	-1719,593	4563,275	1908.279
40	517,082	-904,602	723,369	-1488,542	-1796,068	4668,961	1720.200
Total	22002,271	-4723,526	30050,023	-20962,870	-24395,662	88123,753	90093.990

Table 3.3: Total values per country

Country	W_i	N_i	$\Sigma\theta_i$	$\tilde{W}_i$
USA	17131,631	52801,575	4870,641	22002,271
JAP	8168,569	26075,778	-12892,096	-4723,526
EU	19716,072	60849,327	10333,951	30050,023
CHI	2628,625	9836,434	-23591,495	-20962,870
FSU	1942,436	6403,642	-26338,098	-24395,662
ROW	40506,656	118923,057	47617,097	88123,753
Total	90093,989	274889,813	0,000	90093,989

3.5 Chapter summary and Extensions

In this chapter

- We operate under the assumption that countries cooperate because it makes cooperation beneficial for all countries.
- Flam (2006) arguments can enforce this point.
- We develop transfer schemes for individual rationality for the design of international agreements that achieve global optimality in stock pollutant stochastic control problems. Agreements can in principle be negotiated at each period of time.
- These schemes are also suitable in the context of coalitional rationality. We consider of interest to extend this results to n coalitions.
- The technical complexity of the stochastic algorithm arises because we are not solving Linear Quadratic problems. Thus, suitable aproximate DP solutions have to be implemented.
- Further research could be done if we consider uncertainty about the random perturbation, say the variance of the i.i.d. sequence. We propose to estimate the parameter recursively and to include the estimation in the stochastic control problem.

For each country $i \in J$ and each period $t \in \mathcal{T}$ we have obtained the following solutions stock pollution, emissions and value functions for each model

- Cooperative Model ($P1$): Pareto equilibrium

$$\{s_t^W\} \qquad \{e_{it}^W\} \qquad \{W_i(t, s_{t-1}^W)\}$$

- Non-Cooperative Model ($P2$): Nash equilibrium

$$\{s_t^N\} \qquad \{e_{it}^N\} \qquad \{N_i(t, s_{t-1}^N)\}$$

- Model with Monetary transfers ($P3$): Fall-back non-cooperative equilibrium

$$\{s_t^V\} \qquad \{e_{it}^V\} \qquad \{V_i(t, s_{t-1}^V)\}$$

We find of interest to consider a stochastic model with probability performance criteria and to search the existence of an optimal policy. In absence of international cooperation, these optimal policies obtained under this new perspective could be an alternative behavior for each country which finally will help reducing the international stock pollutant. Note that the target value (X) could be chosen by each country, according some particular negotiation. Usually the target value (X) should be a quantity ranging between the marginal non-cooperative value function and the marginal cooperative value function. These schemes are also suitable in the context of coalitional rationality.

We find of interest to consider stochastic performance criteria based on bounds of probability, i.e., MDP with percentile performance criteria where the decision-maker wants to find a policy that achieves a specific value (target) at a specified probability level α.

Further research could be done if we consider uncertainty about the random perturbation (the variance of the i.i.d. sequence). We propose to estimate the parameter recursively and to include the estimation in the stochastic control problem.

Capítulo 4

Controlling the International Stock Pollutant with Policies Depending on Target Values

"Es justamente la posibilidad de realizar un sueño lo que hace que la vida sea interesante."

Paulo Coelho

Abstract

In this chapter a stochastic dynamic game formulation of the economics of international environmental agreements of the stock pollutant control is provided. We consider Cooperative versus Non-cooperative Stochastic Dynamic Games formulated as Markov Decision Processes (MDP). To improve the non-cooperative equilibrium among countries we propose a new alternative where the decision-maker wants to maximize the probability that some total performance of the dynamical game does not exceed a target value during a fixed period of time. The task requirements are therefore formulated as probabilities rather than expectations. This approach is different from the standard MDP, which uses performance criteria based on the expected value of some index. We present properties of the optimal policies obtained under this new perspective. To illustrate our approach a real data based scenario is analyzed.

4.1 Introduction

In the last years, the theory on international environmental agreements (IEA or often called *coalitions*) and the prospect of climate change has motivated many game theoretic studies due to the international dimension of the global environmental resources and the transboundary effects of the polluting activities, i.e. see Fuentes-Albero and Rubio (2010) for an interesting overview.

As Carbone J. C. and Rutherford (2009) pointed out, the current negotiations of a Post-Kyoto agreement suggest that for most countries national self-interest constitutes a dominant guiding principle. This turn is confirmed by a survey among climate policy experts, who anticipate little reductions in global emissions for the year 2020 (see Bohringer and Loschel (2005)). Therefore, there is little hope at the moment that countries will adopt cooperative strategies, even within an international system of tradable emission because countries face a number of incentives in choosing their endowment of emission rights and it is not obvious that permit trade will result in an emission reductions (see Jaehn and Letmathe (2010)).

Although there is a substantial literature in OR that uses game-theoretic and optimal control concepts to analyze stock pollutant control, there are only a few attempts modeling that issue in a stochastic control framework; see for example Haurie and Malhamé (2008), Kolstad and Ulph (2008), Bahn et al. (2008), Fuentes-Albero and Rubio (2010) and references therein.

Our approach differs substantially to the previous models in the literature in the selection of the optimality criteria. The usual ones are average costs and expected discounted costs. Bearing in mind that under mild assumptions these two criteria are equivalent (see Hernández-Lerma and Lasserre (1996)), it is remarkable that our approach differs from the statistical basis shared by these two optimality criteria. We propose to skip from the moment-based optimality to the probability-based optimality. Although optimal policies under the standard criteria (expectation optimality criteria) are computationally simple and useful for many real-life problems, the optimal policies obtained are not reliable when considering a sample or a few samples of a decision process, since only the average performance over many trials is guaranteed to be optimal.

In this chapter, as an extension of the deterministic formulation in Germain et al. (2003), a stochastic theoretical framework for the stock pollutant control is constructed. With our methodology, the cooperative and non-cooperative models are formulated as Markov Decision Processes (MDP) with constraints. Therefore, our approach considers that the decision-maker (country) wants to maximize the probability that some performance of the dynamical system (cost) does not exceed a target value during a fixed period of time $[0, T]$. It follows that the optimization problem to be solved is essentially different from the classical MDP model, see e.g. Altman (1999), Krass and Vrieze (2002) and Puterman (2005). The main results obtained in the paper are: (i) the optimal value functions are distribution functions of some target value fixed by the decision-maker, and (ii) there exists an optimal deterministic Markov policy for the probabilistic optimality criteria proposed problem. Probabilistic criteria for MDP have been considered by Filar and Petrosjan (2000) and Boda et al. (2004). Percentile optimization for MDP are used by Delage and Mannor (2009).

Our approach also differs to the previous attempts in the stochastic games literature because instead of an optimal policy we provide in fact a family of optimal policies for each player. Moreover, in practice the target value could be fixed by the country according to its marginal costs under cooperative and non-cooperative scenarios. From here, in absence of an international optimum, the optimal policies obtained under this new perspective, could drive the alternative behavior of the country to be considered as a potential element of negotiation for reducing the international stock pollutant.

This chapter is organized as follows: In Section 4.2, we report our new proposal of non cooperative model, based on optimization functions with policies depending on target values, with necessary definitions. In Section 4.3, we present an algorithm which computes optimal value functions, optimal action sets, and optimal policies for a finite horizon model. In Section 4.4, we show some numerical results obtained of solver the probabilistic problem, the first applied to an illustrative example with linear reward functions, and the second example based on real scenarios borrowed from Eyckmans and Tulkens (2003) and Casas and Romera (2009a) are presented. Section 4.5 summarizes conclusions and extensions of this chapter.

4.2 Probability Criteria Model

To improve the non-cooperative equilibrium we propose a different criteria to the minimization of the expected total discounted cost. For each country $i \in J$ we consider the problem to find a policy (emission level) which maximize the probability that the discounted total cost does not exceed a specified value X, named target. That is, for the finite horizon model $1 \leq t \leq T$, find a policy $\pi^{i*} = \{e_{it}^*\}$, in the sequel $\pi^{i*} = \pi^*$ (we omit the index i), such that

$$
\begin{aligned}
(P4) \qquad & \max_{\{e_{it}\}_{t\in\mathcal{T}}} \quad \left\{ Prob\left[\left(\sum_{t=1}^{T} \beta^t r_{it}(e_{it}, s_{t-1}) \right) \leq X \right] \right\} \\
& \text{s.t.} \quad s_t = (1-\delta)s_{t-1} + \sum_{i=1}^{n} e_{it} + \xi_t \\
& \qquad e_{it} \geq 0, \quad \forall i = 1, \cdots, n; \quad \forall t = 1, \cdots, T, \\
& \qquad s_0 > 0,
\end{aligned}
$$

where $r_i(e_{it}, s_t) = c_i(e_{it}) + d_i(s_t)$ is a function that measures in monetary terms the total cost incurred by country $i \in J$ from limiting its emissions to e_{it}, and the damages caused by the stock of pollutant s_t during the time period t for the i-th country; $r_{it} \in R$, where R is the finite cost set and R is a subset of $\mathbb{R}$, is a differentiable and convex function ($r_i'' \geq 0$). The functions $c_i(e_{it})$ and $d_i(s_t)$, for all $i \in J$ for all $1 \leq t \leq T$, are the cost function defined in Chapter 2.

Definition 4.1. *The **target value** (X) is any quantity ranging between the expected discounted non cooperative total cost, N_i in (2.8) and the expected discounted cooperative total cost W_i in (2.6).*

Note that problem $(P4)$, given the restrictions, is equivalent to formulations

$$
\min_{\{e_{it}\}_{t\in\mathcal{T}}} \left\{ Prob\left[\left(\sum_{t=1}^{T} \beta^t r_{it}(e_{it}, s_{t-1}) \right) > X \right] \right\} \tag{4.1}
$$

$$
\text{and} \max_{\{e_{it}\}_{t\in\mathcal{T}}} \left\{ \mathbb{E}\left[\mathbf{1}_{R_i \leq X} \right] \right\}, \tag{4.2}
$$

$$
\text{where the discounted total cost is } R_i = \left(\sum_{t=1}^{T} \beta^t r_{it}(e_{it}, s_{t-1}) \right).
$$

In our formulation, when making a decision and taking an action at each state s, the decision maker considers not only the original state but also his updated *target* x. Thus, we consider in fact an expanded model of MDP Γ, defined in 2.1, by enlarging the state space.

A new *hybrid state* $(s, x) \in S \times \mathbb{R}$ is introduced. We refer (s, x) as the *hybrid state* of the decision maker to distinguish it from the system's state s. The dynamic of the system is now as follows: if the initial state of the decision maker is (s,x) and an action e is taken according to (2.5), the decision maker's new *hybrid state* transits from (s,x) to $(j, \frac{x-r}{\beta})$ with probability p^e_{sjr}.

Thus, the extended MDP $\tilde{\Gamma}$ has the following structure

$$\tilde{\Gamma} = (\tilde{S}, E, R, P, \beta), \tag{4.3}$$

$$\text{where } \tilde{S} = S \times \mathbb{R} \quad , \quad E = \bigcup_{(s,x)\in\tilde{S}} E(s,x) = \bigcup_{s\in S} E(s).$$

Note that $E(s,x) = E(s)$, with $(s,x) \in \tilde{S}$. The extended stationary conditional transition probabilities are simply

$$p^e_{s,j,r} := \quad Prob\left(\tilde{s}_{t+1} = \left(j, \frac{x-r}{\beta}\right) / \tilde{s}_t = (s,x), e_t = e\right),$$
$$\forall s, j \in S \quad , \quad e \in E(s) \quad , \quad r \in R \quad , \quad x \in \mathbb{R}.$$

R and β are the same as in the MDP Γ, given by (2.1).

Remark that the target x is important when making decisions and consequently we must define policies which depend both on the state and the target, that is on the hybrid state $\tilde{s} = (s,x)$.

4.2.1 Main results

In this subsection we introduce the definition of policy considered in the paper. The policies discussed depend in fact on target values. The main results obtained are: (i) the optimal value function are distribution functions of the target, and (ii) there exists an optimal deterministic Markov policy.

Definition 4.2. *A **decision rule** π_t at stage t is a conditional transition probability measure on the set of admissible actions $E(s_t)$ given the past history $(\tilde{s}_1, e_1, \cdots, \tilde{s}_{t-1}, e_{t-1}, \tilde{s}_t)$.*

Definition 4.3. *A **policy** π is a sequence of decision rules $\pi = \{\pi_t\}_{t\geq 1}$. The set of all policies is denoted by Π. A policy $\pi = \{\pi_t\}_{t\geq 1} \in \Pi$ is said to be the following.*

***Markov policy**, if each decision rule π_t only depends on the current state at stage t. Moreover π is a Markov policy if each π_t verifies that $\pi_t(\cdot/\tilde{s}_1, e_1, \cdots, \tilde{s}_{t-1}, e_{t-1}, \tilde{s}_t) = \pi_t(\cdot/\tilde{s}_t)$. The set of all Markov policies is denoted by Π_m.*

Stationary policy*, if the policy π is a Markov policy, and the decision rules of π are all identical, that is, $\pi_t = \pi_1, \quad \forall t > 1$ which is denoted by $\pi = \pi_1^\infty$. The set of all stationary policies is denoted by Π_s.*

Deterministic policy*, is any policy π such that all of its decision rules π_t are deterministic. The set of all deterministic Markov and deterministic stationary policies are denoted by Π_m^d and Π_s^d, respectively.*

Note that a transition law P and a policy π determine the conditional probability measure P_π. Let R_τ^π denote the random variable that is the sum of discounted costs generated by policy π for the τ-stage finite horizon problem. That is

$$R_\tau^\pi = \sum_{t=1}^{\tau} \beta^{t-1} r_{it}(e_{it}, s_{t-1}), \qquad \forall \tau \geq 1. \tag{4.4}$$

Bearing in mind formulation (4.1) of ($P4$), let consider the following *objective function* generated by policy $\pi \in \Pi$

$$F_\tau^\pi(s,x) = P_\pi\left(R_\tau^\pi > x / \tilde{s}_1 = (s,x)\right), \quad \forall \tau \geq 1. \tag{4.5}$$

Definition 4.4. *The* ***optimal value functions*** *are given by*

$$F_\tau^*(s,x) = \inf_{\pi \in \Pi} \left\{ F_\tau^\pi(s,x) \right\}, \quad \forall (s,x) \in \tilde{S}, \tau \geq 1.$$

Obviously
$$F_\tau^*(s,x) = \begin{cases} 1 & if \quad x \geq \frac{d(1-\beta^\tau)}{1-\beta} \\ 0 & if \quad x < \frac{b(1-\beta^\tau)}{1-\beta} \end{cases} \tag{4.6}$$

where we define the lower b and the upper d bounds on the costs by $b = \inf\{r : r \in R\}$ and $d = \sup\{r : r \in R\}$, respectively.

Remark: For any policy π independent of targets $\{x_\tau\}_{\tau \geq 1}$, $F_\tau^\pi(s,x)$ is a distribution function of x, but this result does not hold for general policy $\pi \in \Pi$.

In the next step, we introduce some notation necessary to check that dynamic programming operators possess the usual monotonicity properties, e.g. see Puterman (2005).

Let $\mathcal{D} = \{u / u : \tilde{S} \to [0;1], measurable\}$ be the space of measurable functions on the extended decision maker's state space $\tilde{S}$. For any policy π stationary and $u \in \mathcal{D}$, we define the dynamic programming operators: the average cost G with emissions policies e, the average cost over all emissions policies K^π and the minimum of average cost K, by

$$Gu(s,x,e) = \sum_{j\in S, r\in R} p^e_{sjr} u\left(j, \frac{x-r}{\beta}\right); \quad (s,x)\in\tilde{S}, e\in E(s). \tag{4.7}$$

$$K^\pi u(s,x) = \sum_{e\in E(s)} \pi(e/s,x) Gu(s,x,e); \quad (s,x)\in\tilde{S}. \tag{4.8}$$

$$Ku(s,x) = \min_{e\in E(s)}\{Gu(s,x,e)\}; \quad (s,x)\in\tilde{S}. \tag{4.9}$$

Note that

$$(K^\pi)^0 u = u, \quad (K^\pi)^\tau u = K^\pi\left((K^\pi)^{\tau-1}u\right), \quad K^0 u = u, \quad K^\tau u = K(K^{\tau-1}u).$$

Obviously when π is deterministic stationary policy, that is a non-randomized policy such that $\pi_t(\cdot/\tilde{s}_t) = \pi(\tilde{s}_t) \in E(s_t)$, the decision rule π_t at each $t > 1$, is non-random and the policy is a sequence $\pi = (\pi(\tilde{s}_1), \pi(\tilde{s}_2), \cdots)$, we have $K^\pi u(s,x) = Gu\left(s,x,\pi(s,x)\right)$.

In addition, if $r_0 = 0$ and $F^\pi_0(s,x) = P_\pi\left(r_0 \le x, \tilde{s}_1 = (s,x)\right)$ for any policy $\pi = (\pi_\tau, \tau \ge 1)$, then we have

$$F^*_0(s,x) = I_{[0,\infty)}(x) \quad , \qquad \forall (s,x)\in\tilde{S}, \pi\in\Pi, \tag{4.10}$$

where $I_{[0,\infty)}$ is the indicator function of set $[0,\infty)$. We have the following Lemma, which checked that the operators G, K^π and K defined above possess the usual monotonicity properties of dynamic programming.

Lemma 4.1.

(i) If $u, v \in \mathcal{D}$, $u \le v$ then $Gu \le Gv$, $K^\pi u \le K^\pi v$, $Ku \le Kv$.

(ii) Let $u \in \mathcal{D}$. If $u(s,x)$ is a non-decreasing and a left continuous function of x for any $s \in S$, then $Ku(s,x)$ is also non-decreasing and a left continuous function of x for each $s \in S$.

(iii) There exists a deterministic stationary policy f such that $K^f u = Ku$.

Proof.

(i) For each $(s,x) \in \tilde{S}$, $\quad e \in E(s)$,

$$\begin{aligned}
\text{since} \quad u\left(j, \frac{x-r}{\beta}\right) &\leq v\left(j, \frac{x-r}{\beta}\right) \quad , \quad \forall (j,x) \in \tilde{S}, \\
\text{then} \quad \sum_{j\in S, r\in R} p^e_{sjr} u\left(j, \frac{x-r}{\beta}\right) &\leq \sum_{j\in S, r\in R} p^e_{sjr} v\left(j, \frac{x-r}{\beta}\right), \\
\text{and} \quad Gu(s,x,e) &\leq Gv(s,x,e).
\end{aligned}$$

Additionally

$$\begin{aligned}
K^\pi u(s,x) &= \sum_{e\in E(s)} \pi(e/s,x) Gu(s,x,e) \leq \sum_{e\in E(s)} \pi(e/s,x) Gv(s,x,e) \\
&= K^\pi v(s,x).
\end{aligned}$$

Finally, for each $(s,x) \in \tilde{S}$

$$\min_{e\in E(s)} \sum_{j\in S, r\in R} p^e_{sjr} u\left(j, \frac{x-r}{\beta}\right) \leq \min_{e\in E(s)} \sum_{j\in S, r\in R} p^e_{sjr} v\left(j, \frac{x-r}{\beta}\right).$$

Thus $Ku(s,x) \leq Kv(s,x)$ provided that the minimum is taken over a finite set $E(s)$.

(ii) For any $s \in S$, $u(s,x)$ is a non-decreasing function of x, that is

$$\lim_{h\to 0} u(s, x-h) = u(s,x).$$

If

$$x_1 \leq x_2 \quad \Rightarrow \quad u(s,x_1) \leq u(s,x_2).$$

From (i)

$$Gu(s,x_1) \leq Gu(s,x_2), \text{ and } \min Gu(s,x_1) \leq \min Gu(s,x_2),$$

then $Ku(s,x_1) \leq Ku(s,x_2)$, and Ku is a non-decreasing function of x.

$$\lim_{h\to 0} Ku(s,x-h) = \lim_{h\to 0} \min_{e\in E(s)} Gu(s,x-h,e) = \min_{e\in E(s)} Gu(s,x,e),$$

and Ku is a left continuous function of x.

(iii) If $\pi = (\pi_1, \pi_2, \cdots)$ is a deterministic admissible policy, then

$$\begin{aligned}
E(s) \equiv e \quad &\Rightarrow \quad \pi(e/s,x) \equiv 1, \\
K^\pi u(s,x) = Gu(s,x,e) &= \min_{e\in E(s)} Gu(s,x,e) = Ku(s,x), \\
\text{and } K^\pi u(s,x) &= Ku(s,x).
\end{aligned}$$

The existence of such deterministic admissible policy is guaranteed by Hernández-Lerma (1989) Proposition D3 p. 130.

□

Below, we establish the "optimality principle" for the target value criterion problem.

Theorem 4.1.

(i) The optimal value function $\{F_\tau^*, \tau \geq 0\}$ *given by Definition (4.4) satisfies the optimality equations* $F_0^* = I_{[0,\infty)}$, $F_\tau^* = KF_{\tau-1}^*$, $\tau \geq 1$.

(ii) For all $\tau \geq 0$ *and* $s \in S$, $F_\tau^*(s,x)$ *is a distribution function of* x.

(iii) For any $\tau \geq 0$, *there exists a deterministic stationary policy* π *such that* $F_\tau^\pi = F_\tau^*$.

Proof. The Theorem will be proved by induction. If $\tau = 0$ then by (4.10), $F_0^*(s,x) = I_{[0,\infty)}(x)$ and Theorem 4.1 holds. Assume that Theorem 4.1 is true for $\tau = k$, then it follows that for any $s \in S$, $F_k^*(s,x)$ is a distribution function of x. Applying (iii) of Lemma 4.1 to $F_\tau^*(s,x)$ implies that there exists a deterministic stationary policy δ from $\tilde{S}$ to E such that $\delta(s,x) \in E(s)$ and $GF_\tau^*(s,x,\delta(s,x)) = KF_\tau^*(s,x)$ for all $(s,x) \in \tilde{S}$. It also follows that there exists a policy $\delta \in \Pi_s^d$ such that $K^\delta F_\tau^* = KF_\tau^*$. Now, we define policy $\pi = (\delta, \sigma) \in \Pi_m^d$. By Lemma 4.1 (iii) and properties of operator K we have

$$F_{k+1}^*(s,x) \leq F_{k+1}^\pi(s,x) = K^\delta F_\tau^\sigma(s,x) = K^\delta F_\tau^*(s,x) = KF_\tau^*(s,x). \tag{4.11}$$

On the other hand, for any $\tilde{\eta} = (\eta_1, \eta) \in \Pi$, by Lemma 4.1 we have

$$F_{k+1}^{\tilde{\eta}}(s,x) = K^{\eta_1} F_k^\eta(s,x) \geq K^{\eta_1} F_k^*(s,x) \geq KF_k^*(s,x).$$

Thus $F_{k+1}^*(s,x) \leq KF_k^*(s,x)$. Associating it with (4.11), we obtain that $KF_k^* = F_{k+1}^* = F_{k+1}^\pi$. Then, by Lemma 4.1 and Definition (4.4), we conclude that (i), (ii) and (iii) follows for $k = k+1$. The proof is concluded. □

Remark: A MDP problem with probabilistic criterion such as $(P4)$ in principle is regarded as a difficult problem. The introduction of the enlarged the decision marker's state space $\tilde{S}$ as we consider here, provides a solution, a deterministic Markov policy, in the same way as in the classical framework when using expectation criteria.

Note that by solving the problem $(P4)$, for each country $i \in J$ we obtain a new strategy, say e_i^*, which is related to behavior that in somehow fits the gap between cooperative and non-cooperative solutions, e_i^W and e_i^N respectively, for each player.

The target value (X) is a crucial element in our setting, and we find that it could be a relevant issue in economic negotiations concerning abatement stock pollutant policies.

4.3 Algorithm

We present an algorithm which computes optimal value functions, optimal action sets, and optimal policies for a finite horizon model with the probability criteria, using the backward recursion algorithm of dynamic programming adapted to apply to our problem.

4.3.1 DP - Algorithm

We assume that S and R are both finite sets and we let $R = \{r_1, r_2, \ldots, r_m\}$ with $r_1 < r_2 < \ldots < r_m$. Then, by the Proposition 1 and the finiteness of S, E and R, we have the following conclusions:

(i) For each $s \in S$ and $\tau \geq 1$, $F_\tau^*(s,x)$ is a step distribution function of x with finite jump points;

(ii) For each $s \in S$ and $\tau \geq 1$, $e_\tau^*(s,x)$ is a set-valued function from $\mathbb{R}$ to $E(s)$ with finite discontinuity points;

(iii) For each $\tau \geq 1$, there exists an τ stages optimal deterministic Markov policy which k-th decision rule has the structure analogous to that of $F_\tau^*(s,x)$ and $e_\tau^*(s,x)$, $1 < k < \tau$.

The following algorithm is just the proof of the earlier conclusions.

By Proposition 1, we have

$$F_0^*(s,x) = I_{[0,\infty)}(x) \quad , \quad \forall (s,x) \in \tilde{S}, \quad \pi \in \Pi, \quad x \geq 0,$$

$$F_\tau^*(s,x) = \min_{e\in E(s)} \left\{ \sum_{j\in S, r\in R} p_{sjr}^e F_{\tau-1}^* \left(j, \frac{x-r}{\beta} \right) \times I_{[0,\infty)}(x-r) \right\}, \tag{4.12}$$

$$\forall s \in S, \quad x \in \mathbb{R}, \quad \tau \geq 1.$$

Then for notational convenience, define

$$\begin{aligned} b_\tau(s,x,e) &\equiv \sum_{j\in S, r\in R} p_{sjr}^e F_{\tau-1}^* \left(j, \frac{x-r}{\beta} \right) \times I_{[0,\infty)}(x-r) \quad , \forall s \in S, e \in E(s), \\ M_\tau(s,x) &= \min_{e\in E(s)} \{b_\tau(s,x,e)\} \quad , \quad \forall s \in S. \end{aligned}$$

With the Proposition 1, Lemma 1 and Definition 3 of optimal value functions, we obtain the following algorithm.

Step 1. Calculate

$$\begin{aligned} b_1(s,r_k,e) &= \sum_{j\in S} \sum_{r\in R, r\leq r_k} p_{sjr}^e, \quad \forall s \in S, e \in E(s), \\ M_1(s,r_k) &= \min_{e\in E(s)} \{b_1(s,r_k,e)\} \quad , \forall s \in S, \\ E_1^*(s,r_k) &= \{e : e \in E(s), b_1(s,r_k,e) = M_1(s,r_k)\} \quad , \quad \forall s, \in S, \end{aligned}$$

and select an action $g_1(s,r_k) \in E_1^*(s,r_k)$, $k = 1,2,\ldots,m-1$, and an arbitrary action $g_1(s,r_m) \in E(s)$. Then by 4.12 and definition of the optimal action sets

$$F_1^*(s,x) = \begin{cases} 0 & if \quad x < r_1 \\ M_1(s,r_k) & if \quad r_k \geq x \geq r_{k+1}, \quad k = 1,\ldots,m-1 \\ 1 & if \quad x \geq r_m \end{cases}$$

$$E_1^*(s,x) = \begin{cases} E(s) & if \quad x < r_1 \quad or \quad x \geq r_m \\ E_1(s,r_k) & if \quad r_k \geq x < r_{k+1}, \quad k = 1,\ldots,m-1. \end{cases}$$

Let

$$g_1(s,x) = \begin{cases} g_1(s,r_m) & if \quad x < r_1 \quad or \quad x \geq r_m \\ g_1(s,r_k) & if \quad r_k \geq x < r_{k+1}, \quad k = 1,\ldots,m-1. \end{cases}$$

Step 2. Assume that F_l^*, E_l^* and g_l have been calculated and all jump points of $F_l^*(i,x)$ $(\forall i \in S)$ with $x_1 < x_2 < \cdots < x_\rho$ are known. Calculate the elements of set

$$\{\beta x_k + r_h | k = 1, 2, \ldots, \rho, \quad h = 1, 2, \ldots, m\}$$

and denote them by $u_1 < u_2 < \cdots < u_L \quad (L \leq m\rho)$, in an ascending order. Then, for any $j \in S$ and $r \in R$, we have

$$F_l^*(j, \frac{x-r}{\beta}) = \begin{cases} 0 & if \quad\quad x < u_1 \\ F_l^*(j, \frac{u_k - r}{\beta}) & if \quad u_k \geq x \geq u_{k+1}, \quad 1 \leq k \leq L \\ 1 & if \quad\quad x \geq u_L. \end{cases} \tag{4.13}$$

If $r_1 > u_L$, then $I_{[0,\infty)}(u_k - r) = 0, \quad k = 1, \ldots, L$ and, hence, from 4.11 $F_{l+1}^*(i, u_k) = 0$ for all k. Or, there exists some t such that $u_{t-1} \leq r_1 \leq u_t$, note that if $r_1 < u_1$ we can simply define $u_0 = r_1$ and take $t = 1$.

Calculate

$$\begin{aligned} b_{l+1}(s, r_1, e) &= \sum_{j \in S} p_{sjr_1}^e F_l^*(j, 0), \quad s \in S, e \in E(s), \\ b_{l+1}(s, u_k, e) &= \sum_{j \in S, r \in R, r \leq u_k} p_{sjr}^e F_l^*(j, \frac{u_k - r}{\beta}), \quad s \in S, e \in E(s), k \geq N, \\ M_{l+1}(s, r_1) &= \min_{e \in E(s)} \{b_{l+1}(s, r_1, e)\}, \quad \forall s \in S, \\ M_{l+1}(s, u_k) &= \min_{e \in E(s)} \{b_{l+1}(s, u_k, E)\}, \quad \forall s \in S, k \geq N, \\ E_{l+1}^*(s, r_1) &= \{e : e \in E(s), b_{l+1}(s, r_1, e) = M_{l+1}(s, r_1)\}, \quad \forall s \in S, \\ E_{l+1}^*(s, u_k) &= \{e : e \in E(s), b_{l+1}(s, u_k, e) = M_{l+1}(s, u_k)\}, \quad \forall s \in S, k \geq N. \end{aligned}$$

Next, select actions $g_{l+1}(s, r_1) \in E_{l+1}^*(s, r_1)$, $g_{l+1}(s, u_k) \in E_{l+1}^*(s, u_k)$, $k = t, \ldots, L-1$, and an arbitrary action $g_{l+1}(s, u_L) \in E(s)$. Then by 4.11, 4.12 and definition of optimal action sets

$$\begin{aligned} F_1^*(s, x) &= \begin{cases} 0 & if \quad\quad x < r_1 \\ M_1(s, r_k) & if \quad r_k \geq x \geq r_{k+1}, \quad k = 1, \ldots, m-1 \\ 1 & if \quad\quad x \geq r_m \end{cases} \\ E_1^*(s, x) &= \begin{cases} E(s) & if \quad\quad x < r_1 \quad or \quad x \geq r_m \\ E_1(s, r_k) & if \quad r_k \geq x < r_{k+1}, \quad k = 1, \ldots, m-1 \end{cases} \end{aligned}$$

Let the decision rule at the next stage be defined by

$$g_{l+1}(s,x) = \begin{cases} g_{l+1}(s,r_1) & if \qquad r_1 < x \le u_N \\ g_{l+1}(s,u_k) & if \quad u_k < x \le u_{k+1}, \quad k = 1,\ldots,m-1 \\ g_{l+1}(s,u_L) & if \qquad x \le r_1 \quad or \quad x > u_L, \end{cases}$$

Step 3. Repeat Step 2 until $l+1 = \tau$, or $\tau = T$.

We construct the optimal value function F^*_τ and an optimal policy $\pi^* = (g_\tau, g_{\tau-1}, \ldots, g_1)^\infty$. In the process, the corresponding optimal action sets

$$E^*_1(s,x), E^*_2(s,x), \ldots, E^*_\tau(s,x)$$

are constructed as well. By Proposition 1, these sets characterize all τ stages optimal policies.

4.3.2 Modified DP - Algorithm

The DP Algorithm can calculate the optimal value functions, optimal policies and action sets accurately, however, it can quickly become computationally prohibitive. At each iteration more and more points (u_k) need to be considered. For a large state space, a large action space and a large reward set this will have drastic consequences. The number of points that need to be considered and thereby the time to do this will grow exponentially.

To overcome this problem a new algorithm is presented below. This algorithm approximates the solution found by the DP Algorithm by calculating a fixed number of points at each iteration. However, by taking this number large enough, the approximation will be quite good and the computational time will decrease significantly. We will that all rewards in the problem are positive.

The idea is that (irrespective of the iteration index l) a bounded monotone decreasing function such as $F^*_l(i,x)$ on an interval $[0, v_m]$ can be well approximated by an array of values

$$\{(v_1, F^*_l(i,v_1)), (v_2, F^*_l(i,v_2)), \cdots, (v_m, F^*_l(i,v_m))\}$$

provided that $|v_{i+1} - v_i|$ is sufficiently small. The interpolation between the values $F^*_l(i,v_i)$ and $F^*_l(i,v_{i+1})$ at v_i and v_{i+1} can be carried out in a number of ways. In the implementation below upper end is used. That is,

$$F^*_l(i,v) = F^*_l(i,v_{i+1}) \qquad \forall v \in (v_i, v_{i+1}]$$

The following enhanced dynamic programming algorithm can now be used. For notational convenience, we assume $\beta = 1$ and define

$$\begin{aligned} b_\tau(i,x,e) &\equiv \sum_{j\in S, r\in R} p^e_{ijr} F^*_{\tau-1}(j, x-r) \times I_{[0,\infty)}(x-r) \qquad \forall i \in S, e \in E(i) \\ M_\tau(i,x) &= \min_{e\in E(i)} \{b_\tau(i,x,e)\} \qquad \forall i \in S \end{aligned}$$

Step 1. Initialize:

Choose m points $v_1 < v_2 < \cdots < v_k < \cdots < v_m$ that will represent the target values. The value of v_1 needs to be N_i^T, the expected discounted total cost of non cooperative problem. The value of v_m is the largest target value that will be computed. The larger the m, the more accurate the approximation of the optimal value functions will be. Taking equi-spaced v_k's will have computational advantages. Now by Proposition 1

$$F^*_0(i,x) = \begin{cases} 0 & if \qquad x \le v_1 \\ 1 & if \quad v_{k-1} < x \le v_k, \quad k = 2, \ldots, m. \end{cases}$$

Step 2. Assume that F^*_l has been calculated. Now calculate

$$\begin{aligned} b_{l+1}(i,v_k,e) &= \sum_{j\in S, r\in R} p^e_{ijr} F^*_l(j, v_k - r), \qquad i \in S, e \in E(i), k \ge 1, \\ M_{l+1}(i,v_k) &= \min_{e\in E(i)} \{b_{l+1}(i,v_k,e)\}, \qquad \forall i \in S, k \ge 1, \\ E^*_{l+1}(i,v_k) &= \{e : e \in E(i), b_{l+1}(i,v_k,e) = M_{l+1}(i,v_k)\}, \qquad \forall i \in S, k \ge 1. \end{aligned}$$

Next, select actions $g_{l+1}(i,v_k) \in E^*_{l+1}(i,v_k)$, with $k = 1, \ldots, m$. Then

$$F^*_{l+1}(i,x) = \begin{cases} 0 & if \qquad x \le v_1 \\ M_{l+1}(i,v_k) & if \quad r_{k-1} < x \le v_k, \quad k = 2, \ldots, m \end{cases}$$

$$E^*_{l+1}(i,x) = \begin{cases} E(i) & if \qquad x \le v_1 \\ E^*_{l+1}(i,v_k) & if \quad v_{k-1} < x \le v_k, \quad k = 2, \ldots, m. \end{cases}$$

Let the decision rule at the next stage be defined by

$$g_{l+1}(i,x) = \begin{cases} g_{l+1}(i,v_1) & if \qquad x < v_1 \\ g_{l+1}(i,v_k) & if \quad v_{k-1} < x \le v_k, \quad k = 2, \ldots, m. \end{cases}$$

Step 3. Repeat Step 2 until $l + 1 = \tau$ or $l + 1 = T$.

The approximate optimal value function F_τ^* and an optimal policy $\pi^* = (g_\tau, g_{\tau-1}, \ldots, g_1)^\infty$ have now been constructed. In the process, the corresponding approximate optimal action sets $E_1^*(i,x), E_2^*(i,x), \ldots, E_\tau^*(i,x)$ have been constructed as well. By Proposition 1, these sets characterize all τ stages optimal policies.

4.4 Numerical Results

In this section, we show some numerical results obtained of solver the probabilistic problem $(P4)$, the first applied to an illustrative example with linear reward functions r_i, and the second example applied to a real scenario. The simulations are made for a time horizon of 100 years, we give the results only up to 2030, in order to avoid boundary problems. All computations were made by use of the software Matlab 7.3.0 (R2006b).

We have implemented the equivalent formulation of problem $(P4)$ given in Section 3, following the objective function equivalents (4.1) and (4.2). Thus, we have developed specific code for our example.

4.4.1 The Linear Case

One assumes, by definition of reward function, that

$$r_i(e_{it}, s_t) = c_i(e_{it}) + d_i(s_t)$$

where $c_i(e_{it}) = ae_{it}$, and $d_i(s_t) = bs_t + c$, $\forall a, b, c \in \mathbb{R}$.

Following (4.1) and (4.2) the objective linear function of the probabilistic model $(P4)$ has the following form

$$\max_{\{e_{it}\}_{t\in\mathcal{T}}} \quad \left\{ Prob\left[\left(\sum_{t=1}^{T} \beta^t \left(c_i(e_{it}) + d_i(s_t)\right)\right) \le X\right]\right\}$$

Then the objective function of model $(P4)$ is equal to the objective function of the following linear programming

$$\max_{\{e_{it}\}_{t\in\mathcal{T}}} \quad \left\{ Prob\left[\left(\sum_{t=1}^{T} \beta^t \left(ae_{it} + bs_t + c\right)\right) \le X\right]\right\}$$

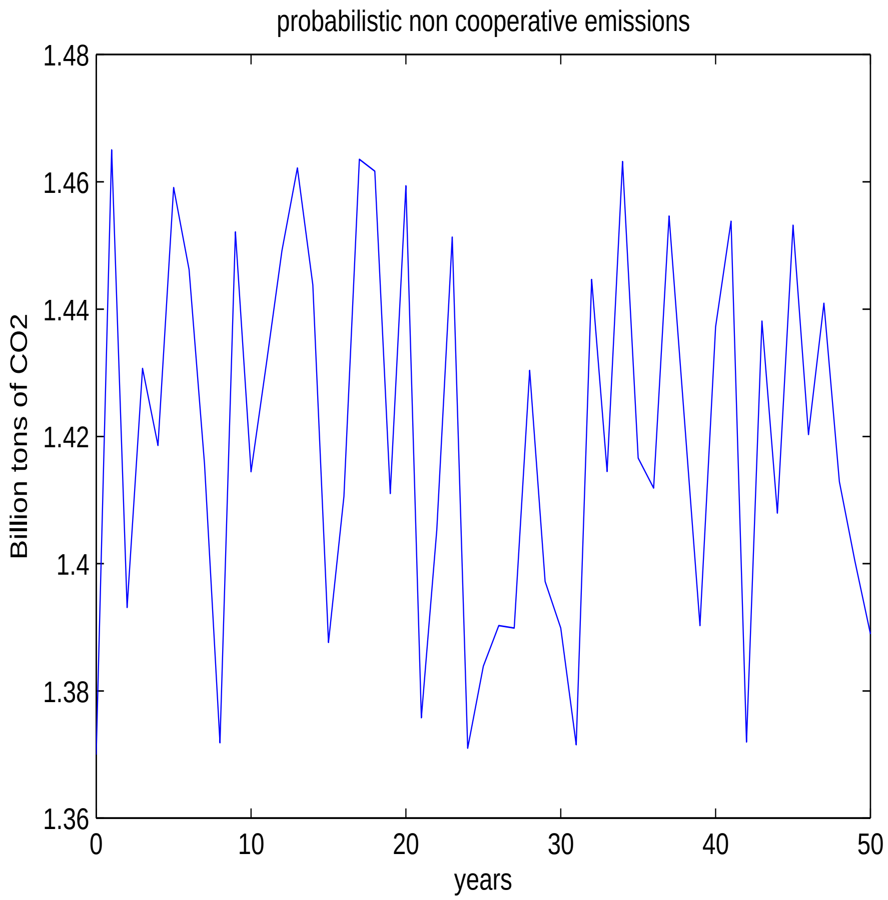

Figure 4.1: Optimal probabilistic non cooperative emissions e_{it}^{l} for country i at each period of time t in billion tons of carbon equivalent.

We use as target value $X = 220000$, and the initial conditions following

$$e_{1990} = [1.37, \quad 0.29, \quad 0.872, \quad 0.805, \quad 1.066, \quad 3.43]$$

the initial CO_2 vector of emissions e_{1990}, in absence of any control are taken from the RICE model and these emissions are measured in billion tons of carbon, the initial stock or preindustrial level of the CO_2 atmospheric stock, is taken as 590 billion tons of carbon equivalent ($s_0 = 590$). Finally the discount factor per year, that appears in the objective function of problem $(P4)$ is taken as

$$\beta = \frac{1}{(1+\rho)^1} = 0.98$$

where the annual discount rate is chosen as $\rho = 0.02$.

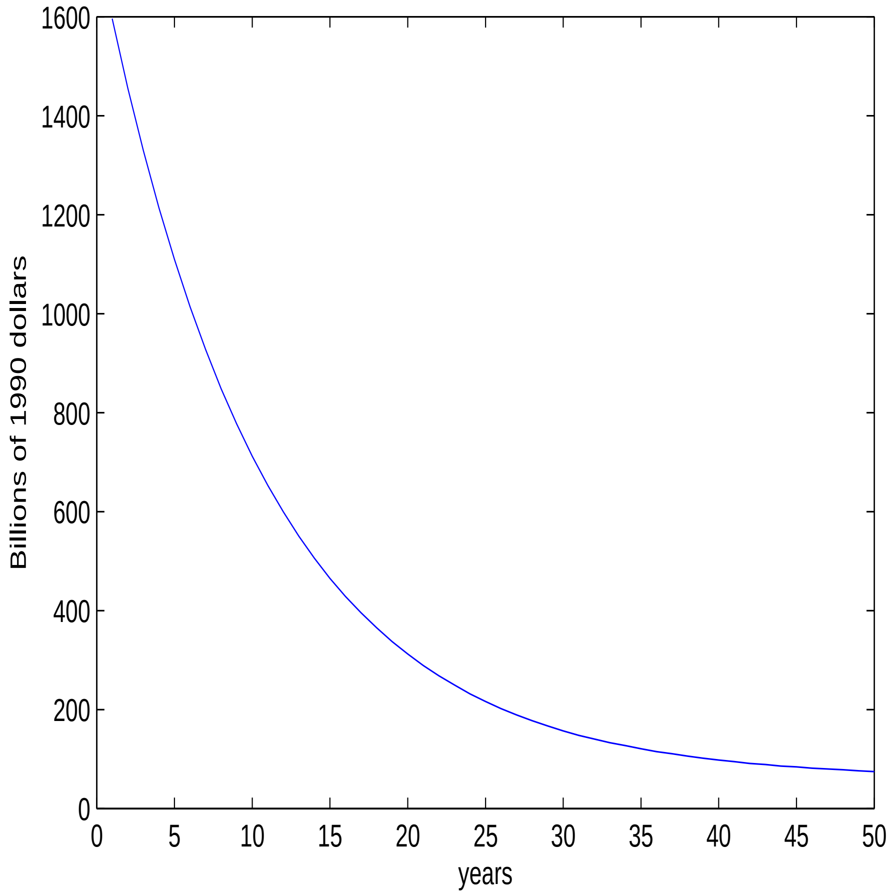

Figure 4.2: Optimal Cooperative Value Function P_{it}^{l} per country i for each period of time t in billions of 1990 USA dollars.

The random disturbance ξ_t is a noise process as in (2.3), i.e. sequence of i.i.d. random variables and independent of the initial state s_0, with normal distribution and

$$\mathbb{E}\left[\xi_t\right] = 0, \qquad \sigma^2 = \mathbb{E}\left[\xi_t^2\right] = 10, \qquad \forall t = 1, 2, ..., T-1.$$

In our simulations we have estimate the expectation of the damages functions and its probability correspondences, over 1000 runs carried out after the corresponding 1000 values of the standard normal disturbance ξ_t. We obtain the probability value of $Prob = 0.8150$ with the target value equal to 220000.

Table 4.1 column 1 gives the optimal cooperative emissions e_{it}^{*} in billion tons of CO_2 equivalent for each country during each period of time t. These results are related with problem $(P4)$. The last row gives the cumulated emissions per country until the end of the

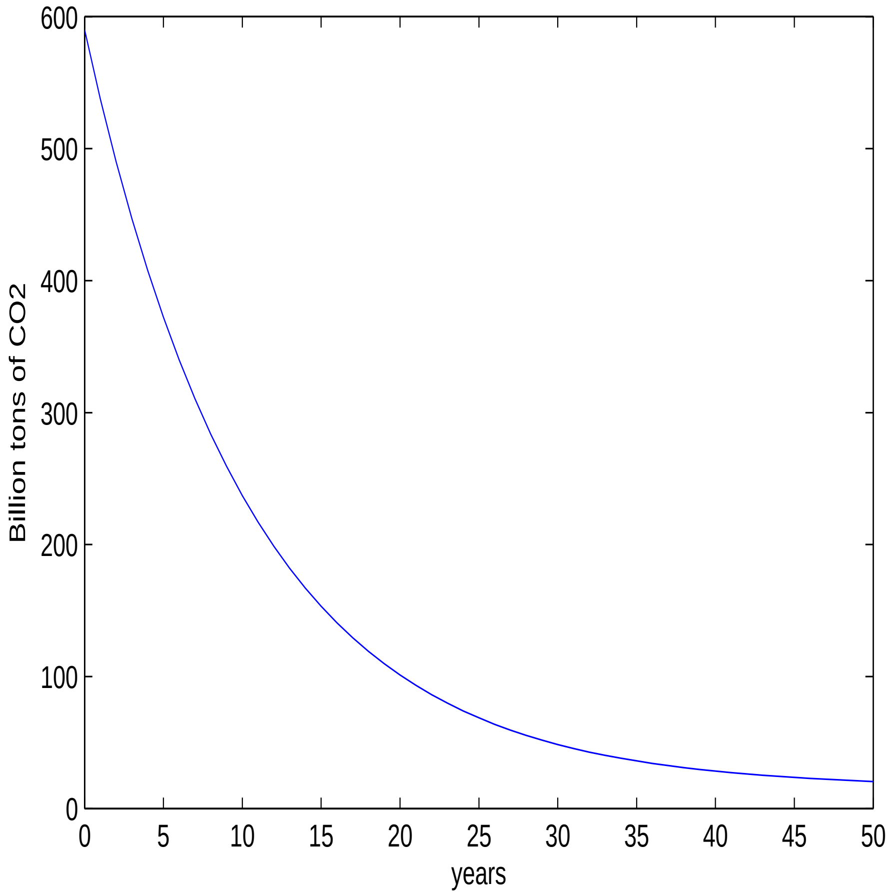

Figure 4.3: Optimal probabilistic non-cooperative stocks for the linear case

horizon T in billion tons of carbon. Figure 4.4.1 shows the optimal cooperative emissions e^*_{it} for country i and per each period of time t.

Table 4.1 column 2 gives the optimal probabilistic value function P^l_{it} for the country i during each period of time t in billions of 1990 USA dollars. These results are related with problem $(P4)$. The last row gives the cumulated value function per country and the total of the world at the end of the final period T, measured in billions of 1990 USA dollars. Figure 4.4.1 shows the optimal cooperative value function P^l_{it} for the country i and per each period of time t in billions of 1990 USA dollars.

Table 4.1 column 3 gives the probabilistic non cooperative optimal stock of pollutant, s^l_t at each period of time t in billion tons of carbon.

Figure 4.4.1 depicts the optimal probabilistic and non-cooperative stocks of pollutant, s_t^l for each period of time t in billion tons of carbon equivalent.

Table 4.1: Optimal probabilistic non cooperative emissions e^l_{it}, Value Function P^l_{it} and stock s^l_t for country i at each period of time t in billion tons of carbon equivalent for linear case with target X.

t	e^l_{it}	P^l_{it}	s^l_t
1	1.4650	1596.1820	537.8340
2	1.3931	1456.0505	490.3380
3	1.4307	1329.4736	447.1970
4	1.4186	1214.0498	407.9654
5	1.4591	1109.5905	372.3404
6	1.4462	1014.2473	339.9409
7	1.4156	927.3518	310.4559
8	1.3719	848.1165	283.6073
9	1.4521	777.1436	259.2796
10	1.4145	711.7523	237.1255
11	1.4315	652.7073	217.0024
12	1.4492	599.0960	198.7260
13	1.4622	550.3752	182.1240
14	1.4438	505.8222	167.0128
15	1.3876	464.8828	153.2189
16	1.4106	428.2401	140.7019
17	1.4635	395.3043	129.3756
18	1.4617	365.0139	119.0771
19	1.4110	336.9918	109.6640
20	1.4594	312.3067	101.1549
21	1.3758	288.7449	93.3357
22	1.4053	268.1352	86.2568
23	1.4513	249.6661	79.8674
24	1.3710	231.8015	73.9784
25	1.3839	216.1882	68.6377
26	1.3903	201.9763	63.7888
27	1.3899	189.0125	59.3802
28	1.4304	177.6265	55.4130
29	1.3972	166.6980	51.7731
30	1.3899	156.8977	48.4568
31	1.3715	147.8542	45.4236
32	1.4447	140.4641	42.7393
33	1.4145	132.9939	40.2688
34	1.4632	126.8678	38.0716
35	1.4166	120.5386	36.0275
36	1.4119	115.0288	34.1644
37	1.4546	110.4684	32.5135
38	1.4225	105.7413	30.9805
39	1.3903	101.3280	29.5547
40	1.4372	97.9771	28.3054

4.4.2 The Real Scenario Case

In this section, we show some numerical results obtained of solver the probabilistic problem $(P4)$, in a real scenario. The simulations are made for a time horizon of 100 years, we give the results only up to 2030, in order to avoid boundary problems. All computations were made by use of the software Matlab 7.3.0 (R2006b). We have implemented the equivalent formulation of problem $(P4)$ given in Section 3, following the objective functions (4.1) and (4.2). Thus, we have developed specific code for our example.

Table 4.2: Target values (X) and the optimal value functions (Probability) for country i.

Target(USA)	53000	53500	53800	54000	55000	**55200**
Prob(USA)	0.1090	0.2710	0.4490	0.5340	0.8670	**0.8920**
Target(JAP)	26000	26500	26800	27200	**27300**	27400
Prob(JAP)	0.0541	0.3392	0.5814	0.8590	**0.8930**	0.9380
Target(EU)	60000	61500	63000	63500	**63650**	63700
Prob(EU)	0.0130	0.2231	0.7780	0.8749	**0.8980**	0.9106
Target(CHI)	10000	10200	**10250**	10270	10300	10500
Prob(CHI)	0.3810	0.8440	**0.885**0	0.9080	0.9280	0.9950
Target(FSU)	6500	6650	6675	6677	**6680**	6700
Prob(FSU)	0.2361	0.8310	0.8810	0.8670	**0.9010**	0.9130
Target(ROW)	119000	12000	121000	123000	**124500**	125000
Prob(ROW)	0.0950	0.2071	0.3891	0.7520	**0.9090**	0.9360

The cost and damage functions used in this case are nonlinear and the arguments of these functions are selected according climate and economic principles. The temperature change equation is taken from the climate economy model RICE (Regional Integrated model of Climate and the Economy), as well as most of the parameter values and all basic data on GDP, population, capital stock, carbon emissions and concentration and global mean temperature. A complete overview of the equations and parameter values of the Climate Negotiation (CLIMNEG) World Simulation Model (abbreviated as CWSM) can be found in Eyckmans and Tulkens (2003) and Casas and Romera (2009a). The division of the world is the same as in the RICE model. There are 6 countries or regions: USA, Japan, European Union (EU), China, Former Soviet Union (FSU) and Rest of the World (ROW). The time is divided in years, the initial period (period $t = 0$) refers to year 1990.

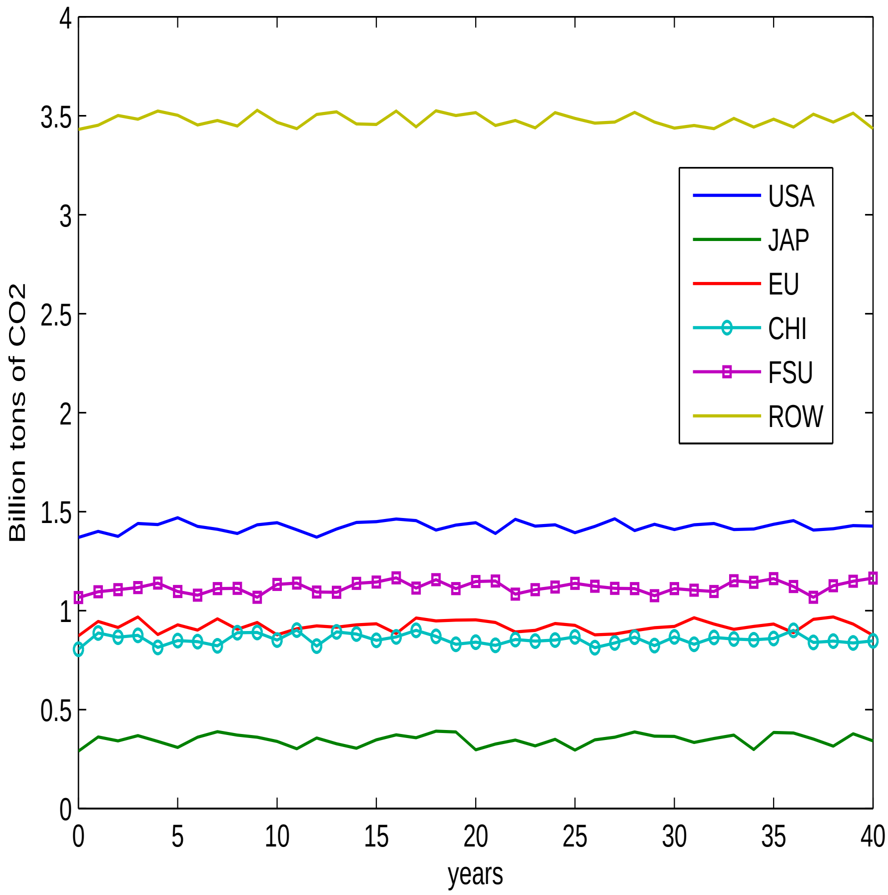

Figure 4.4: Optimal emissions e^*_{it} for country i at each period t with target X.

We have considered several simulation scenarios for the target value X. Table 4.2 shows the outputs we have obtained according to solutions of problem $(P4)$ for different choices of the target values, x_i with $1 \leq i \leq 6$, following Definition 4.1. As an example, we have analyzed more extensively one particular case of the target vector X, that guarantees with a probability close to 0.90 for all countries, that cumulated discounted costs will not be larger that the corresponding target values, x_i with $1 \leq i \leq 6$. More precisely we consider, in billions of 1990 USA dollars, the target values $X = [55200, 27300, 63650, 10250, 6680, 124500]$ (values in bold in Table 4.2).

Solving problem $(P4)$ gives us the optimal solution probability vector

$$Prob = [0.912, 0.918, 0.898, 0.883, 0.892, 0.886].$$

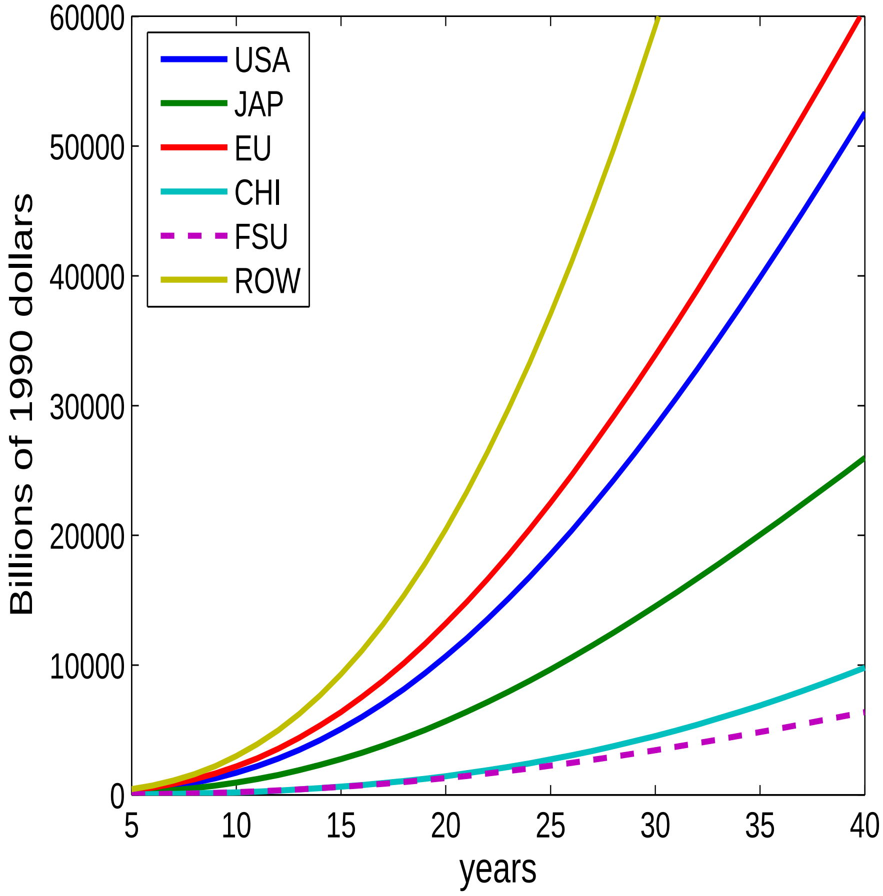

Figure 4.5: Optimal cumulated discounted cost R^*_{it} per country i for each period of time t with target X.

This result can be interpreted as follows. For each country i, $1 \leq i \leq 6$, it can be guarantee that at the corresponding percent, say 91.20 for USA, the cumulated cost incurred by the country adopting the optimal probabilistic emissions, e^*_{it}, will not be larger than its target value, say 55200 billions of 1990 USA dollars.

The optimal emissions, e^*_{it}, in billion tons of CO_2 equivalent, for each country i, per each period of time t given target X, are depicted in Figure 4.4.2. Table 4.3 presents some of these optimal emissions values.

The optimal cumulated discounted costs, R^*_{it}, for each country i during each period of time t in billions of 1990 USA dollars with target X are shown in Figure 4.4.2. Table 4.4 provides some outputs of the values of optimal cumulated discounted costs.

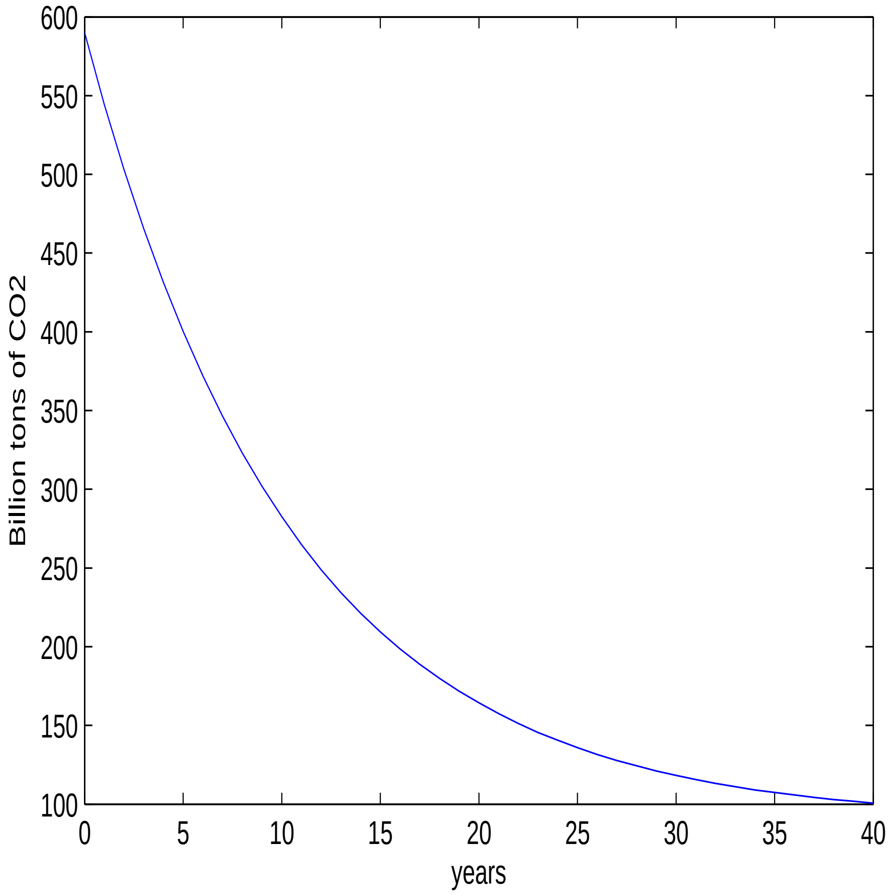

Figure 4.6: Optimal stocks pollutant s_t^* at each period t with target X.

One assumes, by definition of target value, that

$$X_i = [55200, 27300, 63650, 10250, 6680, 124500]$$

where x_i, is the target value for each country i. Each target x_i was found by simulation, country per country. And with this target values, we obtain, for each country and their respective target values, the following optimal probability values

$$Probr = [0.9120, 0.9180, 0.8980, 0.8830, 0.8920, 0.8860]$$

Table 4.3 gives the optimal cooperative emissions e_{it}^r in billion tons of CO_2 equivalent for each country during each period of time t. These results are related with problem $(P4)$. The last row gives the cumulated emissions per country until the end of the horizon T in billion tons of carbon. Figure 4.4.2 shows the optimal cooperative emissions e_{it}^r for country i and per each period of time t with target X.

Table 4.4 gives the optimal probabilistic value function P_{it}^{r} for each country i during each period of time t in billions of 1990 USA dollars with target X. These results are related with problem $(P4)$. The last row gives the cumulated value function per country and the total of the world at the end of the final period T, measured in billions of 1990 USA dollars. Figure 4.4.2 shows the optimal cooperative value function P_{it}^{l} for the country i and per each period of time t in billions of 1990 USA dollars with target X.

Table 4.5 gives the probabilistic non cooperative optimal stock of pollutant, s_t^r at each period of time t in billion tons of carbon with target X.

Figure 4.4.2 depicts the optimal probabilistic and non-cooperative stocks of pollutant, s_t^r for each period of time t in billion tons of carbon equivalent with target X.

The optimal stock of pollutant, s_t^*, at each period of time t in billion tons of carbon equivalent, for each period of time t with target X is depicted in Figure 4.4.2. Table 4.5 provides some outputs of the values of the optimal stock of pollutant.

Table 4.3: Optimal emissions e^*_{it} for each country at each period of time t with target X.

t	USA	Japan	EU	China	FSU	ROW	Total
0	1.3700	0.2920	0.8720	0.8050	1.0660	3.4300	7.8350
1	1.3997	0.3617	0.9453	0.8862	1.0946	3.4522	8.1397
2	1.3749	0.3411	0.9142	0.8660	1.1054	3.5004	8.1021
3	1.4393	0.3676	0.9681	0.8751	1.1163	3.4822	8.2488
4	1.4350	0.3389	0.8792	0.8142	1.1382	3.5233	8.1289
5	1.4683	0.3086	0.9273	0.8475	1.0966	3.5013	8.1497
6	1.4253	0.3601	0.9012	0.8426	1.0772	3.4528	8.0591
7	1.4100	0.3883	0.9578	0.8216	1.1103	3.4750	8.1630
8	1.3899	0.3707	0.9056	0.8883	1.1127	3.4472	8.1143
9	1.4325	0.3604	0.9400	0.8889	1.0675	3.5269	8.2161
10	1.4433	0.3385	0.8773	0.8502	1.1324	3.4656	8.1073
11	1.4076	0.3015	0.9077	0.9007	1.1384	3.4349	8.0907
12	1.3710	0.3565	0.9218	0.8197	1.0942	3.5055	8.0687
13	1.4120	0.3265	0.9154	0.8920	1.0922	3.5195	8.1576
14	1.4454	0.3040	0.9282	0.8819	1.1368	3.4586	8.1550
15	1.4494	0.3467	0.9337	0.8494	1.1444	3.4551	8.1786
16	1.4620	0.3723	0.8833	0.8671	1.1646	3.5233	8.2726
17	1.4545	0.3574	0.9618	0.9002	1.1133	3.4431	8.2303
18	1.4068	0.3899	0.9475	0.8690	1.1563	3.5241	8.2935
19	1.4321	0.3862	0.9511	0.8297	1.1111	3.5002	8.2104
20	1.4431	0.2959	0.9535	0.8403	1.1465	3.5148	8.1940
21	1.3894	0.3260	0.9390	0.8238	1.1489	3.4509	8.0780
22	1.4605	0.3449	0.8921	0.8541	1.0826	3.4755	8.1096
23	1.4269	0.3162	0.8993	0.8459	1.1054	3.4381	8.0318
24	1.4332	0.3497	0.9346	0.8514	1.1181	3.5151	8.2021
25	1.3934	0.2949	0.9257	0.8661	1.1378	3.4862	8.1042
26	1.4249	0.3471	0.8780	0.8121	1.1229	3.4619	8.0469
27	1.4632	0.3601	0.8809	0.8364	1.1121	3.4675	8.1201
28	1.4035	0.3862	0.8991	0.8658	1.1105	3.5168	8.1820
29	1.4356	0.3651	0.9129	0.8225	1.0748	3.4672	8.0780
30	1.4092	0.3640	0.9194	0.8671	1.1103	3.4374	8.1074
31	1.4327	0.3332	0.9629	0.8296	1.1026	3.4500	8.1110
32	1.4399	0.3534	0.9316	0.8637	1.0963	3.4349	8.1199
33	1.4097	0.3703	0.9049	0.8556	1.1512	3.4867	8.1784
34	1.4114	0.2984	0.9198	0.8515	1.1419	3.4422	8.0652
35	1.4355	0.3845	0.9317	0.8591	1.1610	3.4822	8.2541
36	1.4538	0.3816	0.8881	0.8992	1.1218	3.4417	8.1862
37	1.4072	0.3502	0.9549	0.8392	1.0674	3.5070	8.1259
38	1.4125	0.3154	0.9676	0.8452	1.1256	3.4675	8.1338
39	1.4295	0.3773	0.9316	0.8358	1.1476	3.5123	8.2341
40	1.4266	0.3413	0.8749	0.8462	1.1637	3.4347	8.0873
Total	57.0005	13.9325	36.8693	34.2008	44.7516	139.0818	

Table 4.4: Optimal cumulated discounted costs R^*_{it} for each country i for each period of time t in billions of 1990 USA dollars.

t	USA	Japan	EU	China	FSU	ROW	Total
1	4.949	5.366	6.843	0.544	0.605	7.952	26.258
2	19.517	11.779	25.862	1.904	2.373	31.761	93.196
3	43.159	26.372	57.065	4.296	5.242	71.152	207.286
4	75.575	42.815	99.157	7.786	9.170	126.127	360.630
5	116.150	65.099	151.599	12.175	14.107	196.526	555.656
6	164.690	92.268	213.768	17.743	20.044	282.556	791.069
7	220.261	123.562	284.116	24.530	26.683	382.268	1061.420
8	282.720	156.028	362.521	31.667	34.202	496.857	1363.995
9	349.980	191.995	446.924	40.284	42.706	622.404	1694.293
10	423.282	230.876	538.664	50.521	51.195	763.889	2058.427
11	502.264	272.354	634.226	60.595	60.609	917.211	2447.260
12	585.973	315.002	734.297	73.960	70.972	1078.946	2859.150
13	670.194	359.128	837.257	85.197	81.373	1251.438	3284.587
14	757.295	404.690	942.082	99.059	91.512	1435.908	3730.545
15	847.112	449.534	1048.226	114.474	102.285	1626.273	4187.904
16	937.201	494.782	1155.903	129.261	112.955	1818.940	4649.043
17	1029.170	540.184	1259.400	144.330	124.807	2025.260	5123.152
18	1122.824	584.754	1364.138	162.018	135.112	2227.238	5596.084
19	1213.448	629.149	1467.929	181.034	147.083	2440.439	6079.082
20	1304.186	675.464	1569.906	198.810	157.396	2655.098	6560.859
21	1399.364	717.210	1671.883	218.505	168.415	2880.092	7055.469
22	1484.428	758.421	1772.513	236.578	181.007	3100.321	7533.267
23	1575.614	800.894	1868.132	256.985	191.180	3328.057	8020.862
24	1659.194	837.317	1954.929	276.758	201.032	3542.063	8471.293
25	1745.600	878.296	2042.489	296.672	210.554	3766.427	8940.038
26	1824.538	910.501	2130.305	320.548	220.910	3991.662	9398.464
27	1898.579	943.259	2208.315	340.465	230.684	4209.940	9831.241
28	1976.696	973.152	2279.085	360.122	239.658	4419.745	10248.458
29	2045.803	1003.790	2348.330	384.243	249.861	4639.845	10671.872
30	2116.553	1031.572	2412.971	403.657	256.846	4854.243	11075.842
31	2179.239	1059.499	2470.066	427.740	265.139	5061.795	11463.478
32	2240.120	1081.922	2528.721	447.933	273.004	5268.084	11839.784
33	2300.321	1102.261	2581.613	470.361	277.833	5461.790	12194.179
34	2357.107	1130.231	2629.724	493.240	285.203	5667.494	12562.999
35	2403.557	1139.425	2668.665	514.610	290.514	5850.514	12867.285
36	2449.690	1155.875	2711.899	534.613	298.142	6042.108	13192.327
37	2501.945	1173.885	2741.848	560.813	306.571	6224.744	13509.806
38	2545.572	1191.953	2773.867	583.294	309.397	6413.434	13817.516
39	2582.507	1197.174	2804.503	606.212	313.324	6584.189	14087.908
40	2623.725	1212.716	2840.040	629.036	317.748	6774.550	14397.814
Total	52580.100	25970.554	60639.783	9802.571	6377.453	118539.337	273909.797

Table 4.5: Optimal stocks of pollutant s_t^* for each period t in billion tons of carbon equivalent with target X.

t	s_t^r
0	590.0000
1	544.5887
2	503.2917
3	465.7098
4	431.6674
5	400.5691
6	372.1760
7	346.5303
8	323.1425
9	301.9700
10	282.6366
11	265.1027
12	249.1796
13	234.5776
14	221.3860
15	209.4455
16	198.5148
17	188.5776
18	179.5473
19	171.2710
20	163.9237
21	157.1870
22	150.8389
23	145.2260
24	140.0085
25	135.2929
26	131.0180
27	127.2437
28	123.7602
29	120.5500
30	117.7626
31	115.1078
32	112.8148
33	110.6808
34	108.6996
35	106.9226
36	105.3961
37	103.9672
38	102.7126
39	101.6355
40	100.5178

4.5 Chapter summary and Final Remarks

In this chapter we propose a new model with probabilistic performance criteria and we obtain the existence of an optimal policy. In absence of international cooperation, these optimal policies obtained under this new perspective could be an alternative behavior for each country which finally will help reducing the international stock pollutant. Note that the target value (X) could be chosen by each country, according some particular negotiation. Usually the target value (X) should be a quantity ranging between the non-cooperative value function and the cooperative value function. We propose to explore these schemes in the context of coalitional rationality.

We find of interest to consider stochastic performance criteria based on bounds of probability, i.e., MDP with percentile performance criteria where the decision-maker wants to find a policy that achieves a specific value (target) at a specified probability level α, or α-percentile criteria (see Delage and Mannor (2009)).

Further research could be done if we consider uncertainty about the random perturbation (the variance of the i.i.d. sequence). We propose to estimate the parameter recursively and to include the estimation in the stochastic control problem.

It also seems of interest to consider the α-percentile problem under this scenario, say, optimization problem with target values fixed by country.

Capítulo 5

Conclusiones y Extensiones

"Once more unto the breach,
dear friends, once more."

William Shakespeare

En esta tesis se han estudiado cuatro problemas dinámicos y estocásticos, con horizonte discreto y finito, que responden al objetivo de minimizar los daños al medio ambiente que surgen del Nivel (stock) de Contaminaci´on que se acumula en la atmósfera, formulados como Juegos Markovianos. Estos problemas se han modelizado como Procesos de Decisión de Markov con restricciones. A lo largo de toda la tesis se ha utilizado un ejemplo con datos reales, considerando seis países o regiones como jugadores, sus correspondientes emisiones, las funciones de costes y daños generadas por la evolución en el tiempo del stock de contaminación, tomando como datos iniciales el stock y las emisiones de 1990, que son la base de comparación del Protocolo de Kyoto.

Versiones preliminares de algunos de los resultados de contenidos en esta tesis han sido presentados en Conferencias y Seminarios nacionales e internacionales como: X CONFERENCIA DE BIOMETRÍA de Oviedo 2005 (ver en Casas and Romera (2005b)), Seminarios Internacionales Complutenses de Madrid 2005 (ver en Casas and Romera (2005a)), XXX Congreso de la SEIO de Valladolid 2007 (ver en Casas and Romera (2007)), Conference on Sustainable Resource Use and Economic Dynamics SURED 2008 de Ascona, Suiza (ver en Casas and Romera (2008)), Second Workshop on Dynamic Games in Management Science de Valladolid 2009 (ver en Casas and Romera (2009d)) y XII CONFERENCIA DE

BIOMETRÍA de Cadiz 2009 (ver en Casas and Romera (2009c)).

Parcialmente, resultados presentados en esta memoria, contenidos en los documentos de trabajo (working papers) Casas and Romera (2009a) y Casas and Romera (2009b), son los contenidos de los artículos actualmente en revisión en la revista Environmetrics y en la revista European Journal of Operational Research, contenidos que se recogen en los Capítulos 2 y 4 de esta memoria, respectivamente.

Respecto a las extensiones que de forma inmediata pueden darse a los resultados presentados, están en primer lugar la extensión a problemas de horizonte infinito, respecto de los cuales existen, bajo ciertas condiciones para los espacios de estados y de control, así como para los funcionales de coste, resultados teóricos asequibles, como los presentados en Puterman (2005) o en Hernández-Lerma (1989) y en Hernández-Lerma (1999).

Otra línea de extensión inmediata es a funcionales de coste con estructura más compleja, por ejemplo no acotados, para la que existen algunos resultados teóricos en la literatura del área. La dificultad de estas extensiones se centra en la utilización práctica de técnicas de aproximación tipo iteración de valores, para la obtención de óptimos.

El objetivo futuro más retador, desde nuestro punto de vista, es el abordar de forma combinada el problema de estimación y control, mediante la implementación de estimadores recursivos de los parámetros de interés, ver al respecto Romera (2004). En esta línea es también interesante considerar los problemas de control robusto, esto es, la minimización del funcional de coste bajo el peor escenario posible del parámetro (min-max).

Todos estos problemas se encuentran abiertos en escenarios medioambientales, como los considerados en esta tesis.

Apéndice

Transferencias y cooperación con horizon infinito

En el caso de horizonte infinito ($T = \infty$), no es posible utilizar el método recursivo, en este caso podemos considerar la solución estacionaria, teniendo en cuenta que las funciones de costes c_i y de daños d_i no dependen directamente del tiempo. Las formas funcionales de las soluciones sólo varían en el tiempo en función de las variaciones del estado del nivel acumulado de contaminación s.

Los problemas cooperativo $(P1)$ y no cooperativo $(P2)$, se repiten de manera idéntica de período en período. Las funciones solución de estos problemas son ahora constantes, y sólo el valor del stock s deben ser calculado variando con el tiempo.

En el caso donde $T = \infty$, el coste óptimo total descontado esperado del problema cooperativo $(P1)$, no depende ahora explícitamente del tiempo y se puede escribir su función de valor W como:

$$W(s) = \min_{e_i} \mathbb{E}\left[\sum_{i=1}^{n} c_i(e_i) + d_i(\bar{s}) + \beta W(\bar{s})\right]$$

$$\begin{aligned} \text{s.t.} \qquad \bar{s} &= (1-\delta)s + \sum_{i=1}^{n} e_i + \xi \\ e_i &\geq 0 \qquad \forall i \in J \end{aligned}$$

Las condiciones de primer orden, de las ecuaciones de Bellman, conducen a los niveles de emisión y al stock óptimo e^* and s^* respectivamente, ambos en función de s que resuelven el sistema:

$$c_i'(e_i^*) + \sum_{j=1}^{n} d_j'(s^*) + \beta W'(\bar{s}) = 0 \qquad \forall i \in J \tag{5.1}$$

$$s^* = [1-\delta]s + \sum_{i=1}^{n} e_i^* + \xi \tag{5.2}$$

$$W(s) = \sum_{i=1}^{n} c_i(e_i^*) + d_i(s^*) + \beta W(s^*) \tag{5.3}$$

donde c_i' y d_j' son las derivadas de las funciones c_i y d_j respectivamente. Resolviendo estas ecuaciones se obtienen los valores óptimos de emisiones e_i^* y de Stock de polución s^*.

El problema consiste en identificar la función de valor W tal que se verifiquen las condiciones de primer orden (5.1), (5.2) y (5.3). Este sistema se puede considerar como un sistema de ecuaciones funcionales donde las incógnitas son las funciones W, s^* y e_i^* for all $i \in J$. Las incógnónitas de este sistema se pueden calcular siempre que las funciones c_i y d_j sean conocidas para todos $i \in J$.

El mismo razonamiento puede ser aplicado para las funciones de coste por países V_i y W_i. Con horizonte infinito, por similitud con (3.2), (3.3) and ($P3$), el coste total del país i en ausencia de cooperación ahora, sabiendo que cooperará en el futuro, se escribe:

$$V_i(s) = \min_{e_i} \qquad \mathbb{E}\left[(c_i(e_i) + d_i(\bar{s})) + \beta\left[V_i(\bar{s}) + \mu_i[W(\bar{s}) - V(\bar{s})]\right]\right]$$

$$\begin{aligned} \text{s.t.} \qquad & s^V = (1-\delta)s + \sum_{i=1}^{n} e_i^V + \xi \quad s\text{given} \\ & e_i \geq 0 \qquad \forall i \in J \\ & e_j \quad \text{with} \quad j \neq i \quad \text{given} \end{aligned}$$

Con el equilibrio no cooperativo se obtienen los niveles de emisión y de stock e_i^V y s^V respectivamente, ambos en función del stock s, que cumplen las condiciones de primer orden:

$$c_i'(e_i^V) + d_j'(s^V) + \beta\left[V_i'(s^V) + \mu_i[W'(s^V) - V'(s^V)]\right] = 0 \qquad \forall i \in J \tag{5.4}$$

$$s^V = [1-\delta]s + \sum_{i=1}^{n} e_i^V + \xi \tag{5.5}$$

$$V_i(s) = c_i(e_i^V) + d_i(s^V) + \beta\left[V_i(s^V) + \mu_i[W(s^V) - V(s^V)]\right] \qquad \forall i \in J \tag{5.6}$$

donde c_i' y d_j' son las derivadas de las funciones c_i y d_j respectivamente. Resolviendo estas ecuaciones se obtienen los valores del equilibrio no cooperativo con transferencias de las emisiones e_i^V y del Stock de polución s^V.

El problema consiste en identificar la función de valor V_i tales que se verifiquen las condiciones de primer orden (5.4), (5.5) y (5.6). Este sistema, con W conocido y haciendo $V = \sum_{i=1}^n V_i$, se puede considerar como un sistema de ecuaciones funcionales donde las incógnitas son las funciones V_i, V, s^V y e_i^V for all $i \in J$. Las incógnitas de este sistema se pueden calcular siempre que las funciones c_i y d_j sean conocidas para todos $i \in J$.

Si para todo $i \in J$ se cumple que :

$$W_i(s) = c_i(e_i^V) + d_i(s^*) + \beta\tilde{W}_i(s^*) \leq V_i(s^V) \tag{5.7}$$

es decir que el coste en que incurre el país i cuando coopera es inferior al coste que obtendría sin cooperación, lo cual es considerado como racionalidad individual en el sentido de que cada país tiene interés en participar en la cooperación. En el caso contrario, se puede aplicar un razonamiento similar al anteriormente descrito, y proponer la siguiente transferencia:

$$\theta_i(s) = -[W_i(s) - V_i(s)] + \mu_i\,[W(s) - V(s)]$$

donde $\mu_i \in]0;1[$, for all $i \in J$ and $\sum_{i=1}^n \mu_i = 0$. Por construcción de las transferencias se cumple que

$$\sum_{i=1}^n \theta_i(s) = 0.$$

Por otra parte, si el país i recibe $\theta_i(s)$ en caso de cooperación, entonces

$$\tilde{W}_i(s) = W_i(s) + \theta_i(s) = V_i(s) + \mu_i[W(s) - V(s)] \leq V_i(s) \tag{5.8}$$

esto prueba que la cooperación es individualmente racional cualquiera que sea el stock de contaminación heredado s.

Referencias

Eitan Altman. *Constrained Markov Decision Processes.* Stochastic Modeling. Chapman & Hall/CRC, 1999.

R. Amir. *Stochastic games in economics and related fields: an overview. In Neyman and Sorin*, chapter 30. Kluwer, Dordrecht, 2003.

O. Bahn, A. Haurie, and R. Malhamé. A stochastic control model for optimal timing of climate policies. *Automatica*, 44:1545–1558, 2008. doi: 10.1016/j.automatica.2008.03.004.

Megan Bailey, U.R. Sumaila, and Marko Lindroos. Application of game theory to fisheries over three decades. *Fisheries Research*, 102(1-2):1–8, February 2010.

L. Balbus and A.S. Nowak. Construction of nash equilibria in symmetric stochastic games of capital accumulation. *Mathematics Methods of Operations Research*, 60:267–277, 2004.

S. Barrett. *Enviroment and Statecraft: The Strategy of Environmental Treaty-Making.* Oxford Univertity Press, Oxford, 2003.

D.P. Bertsekas. *Dynamic Programming and Optimal Control*, volume I. Athena Scientific, Belmont, Massachusetts, second edition, 2000.

John R. Birge and François Louveaux. *Introduction to Stochastic Programming.* Springer-Verlag, New York, 1997.

Gian I. Bischi, Ahmad Naimzada, and Lucia Sbragia. Oligopoly games with local monopolistic approximation. *Journal of Economic Behavior and Organization*, 62(3):371–388, March 2007.

Keith Blackburn and Giam P. Cipriani. Intergenerational transfers and demographic transition. *Journal of Development Economics*, 78(1):191–214, October 2005.

Kang Boda, Jerzy A. Filar, Yuanlie Lin, and Lieneke Spanjers. Stochastic target hitting time and the problem of early retirement. *IEEE Transactions on Automatic Control*, 49 No. 3 March:409–419, 2004.

C. Bohringer and A. Loschel. Climate policy beyond kyoto: Quo vadis? a computable general equilibrium analysis based on expert judgments. *Kyklos*, 58(4):467–493, 2005.

Giacomo Calzolari and Luca Lambertini. Export restraints in a model of trade with capital accumulation. *Journal of Economic Dynamics and Control*, 31(12):3822–3842, December 2007.

Carsten Helm Carbone J. C. and Thomas F. Rutherford. The case for international emission trade in the absence of cooperative policy. *Journal of Environmental Economics and Management*, 2009. doi: 10.1016/j.jeem.2009.01.001.

C. Carraro, J. Eyckmans, and M. Finus. Optimal transfers and participation decisions in international environmental agreements. *Review of International Organisations*, 1(4):379–396, 2006. doi: 10.1007/s11558-006-0162-5.

Carlo Carraro and JerzyA. Filar. *Control and Game theoretic models of the environment*, volume Annals of the International Society of Dynamic Games. Birkhauser, September 1995.

Omar Casas and Rosario Romera. The international stock pollutant control: A stochastic formulation. In *Seminarios Internacionales Complutenses. Global public goods management: Lessons from Game Theory and Practice for Post-Kyoto negotiations.*, Madrid, November 2005a. URL `http://www.ieg.csic.es/workshop_gpgm/workshop_gpgm_archivos/Omar_slides.pdf`.

Omar Casas and Rosario Romera. Modelo estocástico para el control del stock de contaminación ambiental. In *Actas de X CONFERENCIA DE BIOMETRÍA*, Oviedo, 25-27 Mayo 2005b.

Omar Casas and Rosario Romera. Control del stock internacional de contaminación ambiental: Un enfoque basado en procesos de decisión de markov con criterios probabilísticos. In Comité Organizador del XXX Congreso Nacional SEIO, editor, *Actas del XXX Congreso Nacional SEIO*, Valladolid, September 2007. Sociedad de Estadística e Investigación Operativa. ISBN: 978-84-690-7249-3.

Omar Casas and Rosario Romera. Controlling the international stock pollutant in markov decision process with policies depending on target values. In *Monte Verità Conference on Sustainable Resource Use and Economic Dynamics, SURED 2008*, Ascona, Switzerland, June 2008. URL `http://www.cer.ethz.ch/sured_2008/programme/SURED-08_Casas_Romera.pdf`.

Omar Casas and Rosario Romera. The international stock pollutant control: A stochastic formulation. Working Paper 09-08, Universidad Carlos III de Madrid, Departamento de Estadística, Statistics and Econometrics Series 04., February 2009a. URL `http://hdl.handle.net/10016/3656`.

Omar Casas and Rosario Romera. Controlling the international stock pollutant with policies depending on target values. Working Paper 09-60, Universidad Carlos III de Madrid, Departamento de Estadística, September 2009b. URL `http://hdl.handle.net/10016/5736`. Statistics and Econometrics Series 19.

Omar Casas and Rosario Romera. Control del stock de contaminación ambiental con criterios de probabilidad. In *Actas de XII CONFERENCIA NACIONAL DE BIOMETRÍA*, Cádiz, 23-25 Septiembre 2009c.

Omar Casas and Rosario Romera. Controlling the international stock pollutant with policies depending on target values. In *Second Workshop on Dynamic Games in Management Science*, Valladolid, 28-30 June 2009d.

Roberto Cellini and Luca Lambertini. A differential oligopoly game with differentiated goods and sticky prices. *European Journal of Operational Research*, 176(2):1131–1144, January 2007.

P. Chander and H. Tulkens. A core-theoretic solution for design of cooperative agreements on transfrontier pollution. *International Tax and Public Finance*, 2:279–294, 1995.

P. Chander and H. Tulkens. The core of an economy with multilateral enviromental externalities. *International Journal of Game Theory*, 26:379–401, 1997.

Gian L Clementi and Hugo A. Hopenhayn. A theory of financing constraints and firm dynamics. *The Quaterly Journal of Economics*, 121(1):229–265, February 2006. doi: 10.1162/qjec.2006.121.1.229.

S. Clemhout and H. Wan. Dynamic common-property resources and environmental problems. *Journal of Optimization Theory and Applications*, 46(4):471–481, 1985. doi: 10.1007/BF00939152.

Christian Cordes. A potential limit on competition. *Journal of Bioeconomics*, 10(2):127–144, August 2008.

Ronald B. Davies and Benjamin H. Liebman. Self-protection: Antidumping duties, collusion, and fdi. *Review of International Economics*, 14(5):741–757, November 2006.

W.Davis Dechert and S.I. O'Donnell. The stochastic lake game: A numerical solution. *Journal of Economic Dynamics and Control*, 30(9-10):1569–1587, September-October 2006. doi: 10.1016/j.jedc.2005.10.010. URL `http://www.sciencedirect.com/dx.doi.org/10.1016/j.jedc.2005.10.010http://algol.ssc.wisc.edu/a4.3d`.

Erick Delage and Shie Mannor. Percentile optimization for markov decision processes with parameter uncertainty. *Operations Research*, Articles in Advance:1–11, August 2009. doi: 10.1287/opre.1080.0685.

Prajit Dutta and Roy Radner. Profit maximization and the market selection hypothesis. *The Review of Economic Studies*, 66(4):769–798, October 1999. URL `http://www.jstor.org/stable/2566920`.

Prajit Dutta and Rangarajan K. Sundaram. *The Equilibrium Existence Problem in General Markovian Games. In: Organizations with Incomplete Information: Essays in Economic Analysis. A tribute to Roy Radner*, chapter 5, pages 159–207. Cambridge University Press, Cambridge and New York, 1998.

Prajit K. Dutta. *Strategies and games : theory and practice*. Cambridge, Mass., 2001.

J. Eyckmans and H. Tulkens. Simulating coalitionally stable burden sharing agreements for the climate change problem. *Resource and Energy Economics*, 25:299–327, 2003.

Charles Figuieres. Markov interactions in a class of dynamic games. *Theory and Decision*, 66(1):39–68, January 2009.

J.A. Filar and L.A. Petrosjan. Dynamic cooperative games. *International Game Theory Review*, 2 (1):47–65, 2000.

M. Finus. *Game Theroy and International Environmental Cooperation*. Edwar Elgar, Cheltenham, UK and Northampton, USA, 2001.

Sjur Didrik Flam. Balanced environmental games. *Computer and Operations Research*, 33: 401–408, 2006.

Lara Fontanella, Luigi Ippoliti, and Paquale Valentini. Environmental pollution analysis by dynamic structural equation models. *Environmetrics*, 18:265–283, 2007. doi: 10.1002/env.835.

Cristina Fuentes-Albero and Santiago J. Rubio. Can international environmental cooperation be bought? *European Journal of Operational Research*, 202:255–264, 2010. doi: 10.1016/j.ejor.2009.05.006.

Drew Funderberg and Jean Tirole. Capital as a commitment: Strategic investment to deter mobility. *Journal of Economic Theory*, 31(2):227–250, December 1983.

Steven A. Gabriel, Julio Manik, and Shree Vikas. Computational experience with a large-scale, multi-period, spatial equilibrium model of the north american natural gas system. *Networks and Spatial Economics*, 3(2):97–122, June 2003. doi: 10.1023/A:1023955701522.

Steven A. Gabriel, Jifang Zhuang, and Supat Kiet. A large-scale linear complementarity model of the north american natural gas market. *Energy Economics*, 27:639–665, 2005. doi: 10.1016/j.eneco.2005.03.007.

Gérard Gaudet and Hervé Lohoues. On limits to the use of linear markov strategies in common property natural resource games. *Environmental Modeling and Assessment*, 13 (4):567–574, November 2008. doi: 10.1007/s10666-007-9118-2.

M. Germain, Ph. Toint, and H. Tulkens. Calculéconomique itératif et stratégique pour les negociations internationales sur les pluies acides entre la finlande, la russie et l'estonie. *Annales d'Economie et de Statistique*, 43:101–127, 1996.

M. Germain, Ph. Toint, H. Tulkens, and Aart de Zeeuw. Transfers to sustain dynamic core-theoretic cooperation in international stock pollutant control. *Journal of Economic Dynamics and Control*, 28:79–99, 2003.

Garrettt Hardin. The tragedy of the commons. *Science*, 162:10–15, 1968.

A. Haurie and L.A. Viguier. A stochastic dynamic game of carbon emission trading. *Environmental Modelling and Assessment*, 8:239–248, 2003.

Alain Haurie and Roland Malhamé. Introduction to the special issue on stochastic modelling, control, and robust optimization at the crossroads of engineering, environmental economics, and finance. *Automatica*, 44:1457–1459, 2008.

Onésimo Hernández-Lerma. *Adaptive Markov Control Processes*. Springer-Verlag, New York, 1989.

Onésimo Hernández-Lerma. *Further topics on discrete-time Markov control processes*. Springer, 1999.

Onésimo Hernández-Lerma and Xianping Guo. *Continuous-Time Markov Decision Processes: Theory and Applications*, volume 62 of *Stochastic modelling and applied probability*. Springer, Germany, 2009.

Onésimo Hernández-Lerma and Jean Bernard Lasserre. *Discrete-Time Markov Control Processes: Basic Optimality Criteria*. Springer-Verlag, New York, 1996.

Onésimo Hernández-Lerma and Rosario Romera. Limiting discounted-cost control of partially observable stochastic systems. *SIAM on Control and Optimization*, 40:348–369, 2001.

Onésimo Hernández-Lerma and Rosario Romera. The scalarization approach to multiobjective markov control problems: why does it works? *Applied Mathematics and Optimization*, 50:279–293, 2004.

Robert M. Hunt. When do more patents reduce r and d? *American Economic Review*, 96 (2):87–91, May 2006.

Rustam Ibragimov. A tale of two tails: peadkedness properties in inheritance models of evolutionary theory. *Journal of Evolutionary Economics*, 18(5):597–613, October 2008.

IPCC. *Climate Change 2007: The Physical Science Basis, Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Intergovernmental Panel on Climate Change, Cambridge, U.K., cambridge university press edition, February 2007. URL `http://www.ipcc.ch/`.

Florian Jaehn and Peter Letmathe. The emissions trading paradox. *European Journal of Operational Research*, 202:248–254, 2010. doi: 10.1016/j.ejor.2009.05.007.

S. Jorgensen and G. Zaccour. Incentive equilibrium strategies and welfare allocation in a dynamic game of pollution control. *Automatica*, 37:29–36, 2001.

S. Jorgensen, G. Martín-Herrán, and G. Zaccour. Agreeability and time-consitency in linear-state differential games. *Journal of Optimization Theory and Applications*, 119:49–63, 2003.

S. Jorgensen, Guiomar Martín-Herrán, and G. Zaccour. Sustainability of cooperation over time in linear-quadratic differential games. *International Game Theory Review*, 2004.

A. Kanudia and R. Loulou. Advanced bottom-up modeling for national and regional energy planning in response to climate change. *International Journal of Environment and Pollution*, 12(2-3):191–216, 1997.

A. Kanudia and P.R. Shukla. Modelling of uncertainties and price elastic demands in energy-environment planning for india. *OMEGA*, 26(3):409–423, 1998.

Klaus Keller, Benjamin M. Bolker, and David F. Bradford. Uncertain climate thresholds and optimal economic growth. *Journal of Environmental Economics and Management*, 48:723–741, 2004.

Charles Kolstad and Alistair Ulph. Learning and international environmental agreements. *Climatic Change*, 89:89:125–141, 2008.

Matthias Krakel and Dirk Sliwka. Strategic delegation and mergers in olygopolistic contests. *Journal of Economics and Business*, 58(2):119–136, March-April 2006.

Dimitry Krass and O.J. Vrieze. Achieving target state-action frequencies in multichain average-reward markov decision processes. *Mathematics of Operations Research*, 27 No. 3 August:545–566, 2002.

S. Kverndokk. Coalitions and side payments in international co2 treaties. In Van Ierland, editor, *International Environmental Economics, Theories, Models and Application to Climate Change, International Trade and Acidification*, volume 4 of *Developments in Environmental Economics*, Amsterdam, 1994. Elsevier.

W. B. MacLeod. Optimal contracting with subjetive evaluation. *The American Economic Review*, 93(1):216–240, March 2003. URL `http://www.jstor.org/stable/3132169`.

Kaveh Madani. Game theory and water resources. *Journal of Hydrology*, x:1–14, 2009. doi: 10.1016/j.hydrol.2009.11.045.

G. Martín-Herrán and J.P. Rincón-Zapatero. Efficient markov perfect nash equilibria: Theory and application to dynamic fishery games. *Journal of Economic Dynamics and Control*, 29:1075–1096, 2005.

Akio Matsumoto and Nobuko Serizawa. Strategic trade policy under isoelastic demand and asymmetric production costs. *The Annales of Regional Science*, 41(3):525–543, September 2007.

William J. McCausland. Time reversibility of stationary regular finite state markov chains. *Journal of Econometrics*, 136(1):303–318, January 2007.

Jean-Francois Mertens and T. Parthasarathy. *Stochastic Games and Related Topics: in honor of Professor L.S. Shapley*, chapter Nonzero-sum stochastic games, pages 145–160. Kluwer Academic, Boston, 1991.

Frederic H. Murphy, Michael A. Toman, and Howard J. Weiss. A dynamic nash game model of oil market disruption and strategic stockpiling. *Operations Research*, 37(6):958–971, 1989.

A. Neyman and S. Sorin, editors. *Stochastic Games and Applications*, Dordrecht, 2003. Kluwer.

W. Nordhaus and J. Boyer. *Warming the World: Economic Models of Global Warming*. MIT Press, Cambridge, MA, 2000.

W.D. Nordhaus and Z. Yang. A regional dynamic general-equilibrium model of alternative climate change strategies. *American Economic Review*, 86:741–765, 1996.

Andrzej S. Nowak. On perfect equilibria in stochastic models of growth with intergenerational altruism. *Economic Theory*, 28:73–83, 2006. doi: 10.1007/s00199-004-0609-y.

United Nations ONU. *Kyoto Protocol*. Framework Convention on Climate ChangeUnited Nations, Kyoto, Japan, December 1997. URL `http://unfccc.int/kyoto_protocol/items/2830.php`.

T. Parthasarathy and S. Sinha. Existence of stationary equilibrium strategies in non-zero sum discounted stochastic games with uncountable state space and state-independent transitions. *International Journal of Game Theory*, 18(2):189–194, 1989. doi: 10.1007/BF01268158.

L.A. Petrosjan and G. Zaccour. Time-consistent shapley-value allocation of pollution cost control. *Journal of Economic Dynamics and Control*, 27:381–398, 2003.

PNUMA. *Panorama futuro y recomendaciones, Perspectivas del Medio Ambiente Mundial 2000*. Programa de las Naciones Unidas para el Medio Ambiente, ONU, Londres, earthscan publications, ltd. edition, 1999. URL `http://grid.cr.usgs.gov/geo2000/ov-es.pdf`.

Warren B. Powell. *Approximate Dynamic Programming: Solving the Curses of Dimensionality*. John Wiley and Sons, Hoboken, New Jersey, 2007.

M.L. Puterman. *Markov Decision Processes:Discrete Stochastic Dynamic Programming*. John Wiley and Sons, New Jersey, 2005.

Roy Radner, Roger Myerson, and Eric Maskin. An example of a repeated partnership game with discounting and with uniformly inefficient equilibria. *The Review of Economic Studies*, 53(1):59–69, January 1986. URL `http://www.jstor.org/stable/2297591`.

Michael Reksulak, William F. Shughart, and Robert D. Tollison. Innovation and the opportunity cost of monopoly. *Managerial and Decision Economics*, 29(8):619–627, 2008.

Rosario Romera. Optimality for the linear quadratic non gaussian problem via the asymmetric kalman filter. *International Journal of Adaptive Control and Signal Processing*, 18: 73–80, 2004.

S. Rubio and B. Casino. Self-enforcing international enviromental agreements with a stock pollutant. *Spanish Economic Review*, 7:89–109, 2005.

Gerhard Sorger. A dynamic common property resource problem with amenity value and extraction costs. *International Journal of Economic Theory*, 1(1):3–19, 2005. doi: 10.1111/j.1742-7363.2005.00002.x. URL `http://dx.doi.org/10.1111/j.1742-7363.2005.00002.x`.

Ernst von Weizsacker, Amory B. Lovins, and L.Hunter Lovins. *Factor four: Doubling wealth. Halving resource use*. Londres, earthscan publications, ltd. edition, 1997.

WBCSD. *World Business Council for Sustainable Development , Definitions*. ONU, September 1999. URL `http://www.wbcsd.ch/aboutdfn.htm`.

Gang Wu, Ying Fan, Lan-Cui Liu, and Yiu-Ming Wei. An empirical analysis of the dynamic programming model of stockpile acquisition strategies for china's strategic pretroleum reserve. *Energy Policy*, 36:1470–1478, 2008. doi: 10.1016/j.enpol.2008.01.007.

James V. Zidek, Gavin Shaddick, Jean Meloche, Chris Chatfield, and Rick White. A framework for predicting personal exposures to environmental hazards. *Environronmental and Ecological Statistics*, 14:411–431, 2007. doi: 10.1007/s10651-007-0028-x.

Y así, del mucho leer y del poco dormir,
se le secó el cerebro de manera que
vino a perder el juicio.

Miguel de Cervantes Saavedra

List of principal notation

$\mathcal{T}$	horizon set of period of time
T	number of stages
t, τ	stages, period of time
e_{it}	carbon emissions (gigaton of carbon, GtC)
	the actions or controls tken by country i at stage t
e_{i0}	initial emissions at 1990
E	action space
s_t	atmospheric carbon concentration (gigaton of carbon, GtC)
	the state of the system at stage t
s_0	atmospheric carbon concentration (gigaton of carbon, GtC)
	the initial state at 1990
S	state space
$E(s)$	set of admissible actions at state s
β	discount factor
δ	the pollutant's natural rate of atmospheric absorption of CO_2
ξ	random disturbance
G	Markovian Game
Γ	Markov Decision Process
R	cost set
r_{it}	the cost incurred by country i at period t
c_{it}	the emission abatement cost function by country i at period t
d_{it}	the environmental damage function by country i at period t
$p^e_{i,j,r}$	conditional transition probabilities
W	the expected value function cooperative
N	the expected value function non cooperative
V	the expected value function non cooperative with transfers
Θ	the transfer
$\tilde{W}$	the expected value function cooperative with transfers
$\tilde{\Gamma}$	the extended Markov Decision Process
X	target value
$\tilde{S}$	hybrid state space
(s_t, x)	hybrid state with target x

Abbreviations

a.s.	almost surely.
BAU	business-as-usual.
BR	best response set.
CLIMNEG	Climate Negotiation.
CWSM	Climate Negotiation (CLIMNEG) World Simulation Model.
DP	Dynamic Programming.
e.g.	is an abbreviation for the Latin words exempli gratia, which mean ”for the sake of example”.
EU	European Union.
FSU	Former Soviet Union.
GDP	Gross Domestic Product (Producto Interior Bruto).
GSAM	Gas Systems Analysis Model
i.e.	is an abbreviation of the Latin words id est, which mean ”that is”.
IEA	International Environmental Agreement.
i.i.d.	independent and identically distributed.
MCM	Markov Control Model.
MCP	Markov Control Process.
MDP	Markov Decision Processes.
MG	Markov Game.
MPE	Markov Perfect Equilibrium.
RICE	Regional Integrated model of Climate and the Economy.
ROW	Rest of the World.
TSO	T-stage stochastic optimization problem.
WBCSD	World Business Council for Sustainable Development.
w.r.t.	with respect to.